NPS55-90-13

NAVAL POSTGRADUATE SCHOOL

Monterey, California

DEFENSE DATA NETWORK (DDN) PERFORMANCE ANALYSIS USING PROBABILITY MODELING

DONALD P. GAVER
PATRICIA A. JACOBS

July 1990

Approved for public release; distribution is unlimited.

Prepared for:
se Communications Agency
n, Virginia

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral R. W. West, Jr.
Superintendent

Harrison Shull
Provost

This report was prepared in conjunction with research funded by Defense Communications Agency, Reston, Virginia.

This report was prepared by:

Unclassified

Security Classification of this page

REPORT DOCUMENTATION PAGE

1a Report Security Classification UNCLASSIFIED	1b Restrictive Markings
2a Security Classification Authority	3 Distribution Availability of Report Approved for public release; distribution is unlimited
2b Declassification/Downgrading Schedule	
4 Performing Organization Report Number(s) NPS55-90-13	5 Monitoring Organization Report Number(s)
6a Name of Performing Organization Naval Postgraduate School 6b Office Symbol (*If Applicable*) OR	7a Name of Monitoring Organization Defense Communications Agency
6c Address (*city, state, and ZIP code*) Monterey, CA 93943-5000	7b Address (*city, state, and ZIP code*) Reston, Virginia
8a Name of Funding/Sponsoring Organization Defense Communications Agency 8b Office Symbol (*If Applicable*)	9 Procurement Instrument Identification Number
8c Address (*city, state, and ZIP code*) Reston, Virginia	10 Source of Funding Numbers Program Element Number / Project No / Task No / Work Unit Accession No

11 Title (*Include Security Classification*) Defense Data Network (DDN) Performance Analysis Using Probability Modeling

12 Personal Author(s) Gaver, D. P. and Jacobs, P. A.

13a Type of Report	13b Time Covered	14 Date of Report (*year, month,day*)	15 Page Count
Technical	From To	1990, July	

16 Supplementary Notation The views expressed in this paper are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 Cosati Codes			18 Subject Terms (*continue on reverse if necessary and identify by block number*)
Field	Group	Subgroup	Packet switching, fluid approximation, bistability, queueing, data transfer

19 Abstract (*continue on reverse if necessary and identify by block number*

Probability models are constructed for (i) guiding selection of packet size when transmission errors occur at a given, non-bursty rate, (ii) describing delay at a Packet Switching Node when all traffic sources to the node follow the same re-transmit policy. "Exponential backoff" is modeled, and optimal choice of retransmit intervals is discussed.

20 Distribution/Availability of Abstract [X] unclassified/unlimited [ ] same as report [ ] DTIC users	21 Abstract Security Classification Unclassified
22a Name of Responsible Individual D. P. Gaver	22b Telephone (*Include Area code*) (408) 646-2605 — 22c Office Symbol OR/Gv, OR/Jc

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted — security classification of this page

All other editions are obsolete — Unclassified

DEFENSE DATA NETWORK (DDN) PERFORMANCE ANALYSIS USING PROBABILITY MODELING

D. P. Gaver
P. A. Jacobs

SUMMARY

The Data Defense Network (DDN) is a large packet switching network that services elements of the U.S. Department of Defense (DOD). The emphasis of this report is to perform probabilistic analysis of certain features of the DDN system with a view of enhancing or "optimizing" measures of service such as data base throughput and the waiting times experienced by data-base-transfer customers. The particular questions addressed, and the models constructed, are in response to tasking statements supplied by personnel from the Defense Communications Engineering Center, Reston, Virginia.

In Section 2 models are presented to study the optimal length of a packet subject to transmission errors. When a data transfer is to occur the total collection of bits that comprise the data base is divided into packets, i.e., subcollections of contiguous bits from the data base plus a header carrying address information. Packets from the source node, S, travel through a number of packet switching nodes (PSN's) to the destination node, D. The number of PSN's encountered by a packet during its transit from S to D is called the number of hops. Each bit in the packet is susceptible to corruption by various sources along the way. If any bit in the packet is so corrupted the

entire packet is viewed as useless and must be retransmitted. A large packet is more likely to suffer contamination than is a small one. On the other hand, each packet of any size has about the same number of header bits and so small packets require the transfer of an uneconomically large amount of header information as compared to the "live" bits that are the data to be transferred. The question addressed in Section 2 is how to choose the number b of live bits per packet when a) the successful transmission probability per bit is p, b) the number of header bits per packet is fixed and equal to h and c) the number of hops is fixed at J. Simple approximate formulas are obtained for the packet length which maximizes the transmission rate of live bits.

A message is comprised of a number of packets. The Appendix A contains a revision of a prior report which introduces an initial model to study the effect of the number of packets in a message upon data transfer time. The model incorporates the following features. At an initial instant (t=0) a group of $m \geq 1$ packets comprising a message is sent from a source S to a destination D. The m packets are sent out simultaneously via the network. The first to arrive and find space for the m-packet message occupies a one-packet space and reserves space for m-1 others and sends an acknowledgement back to S. If no acknowledgement is received by S in time δ each packet is retransmitted; this action is repeated until acknowledgement occurs. Such is necessary because packets may be lost due to encountering full buffers along the way, etc. Once the initial reservation is made a time elapses until the remaining m-1 packets of the message (originals or duplicates thereof) reach D: each subsequent packet experiences delays similar to the first, but need not reserve space in D's buffer. The model of the Appendix is used to obtain an expression for the long run rate of transfer of data packets from source S. The number of packets per message which maximizes the transfer rate is investigated numerically. As noted in the Appendix, while model does account for retransmission by the tagged source, the possible effect of increased congestion due to retransmission by other sources is not explicitly modeled.

In Section 3 models are introduced to study the effect on D's buffer of all sources retransmitting at a retransmission interval of length δ. We also model the behavior of one form of congestion control, exponential backoff, a procedure that increases successive time-out intervals possibly from δ to 2δ, 2δ to 4δ, etc. The models have as inputs the arrival rate of new or original packets to D's buffer, λ; the length of a packet in terms of its deterministic service time s; the retransmission rate $\upsilon = 1/\delta$; and the size of D's buffer, B. The models are of the amount of work in D's buffer. For a moderate original traffic intensity $\lambda s \approx 0.5$ the models indicate the following. For small retransmission rate υ, the amount of work in D's buffer tends to be small. For a large retransmission rate υ, the buffer tends to be nearly full. If the retransmission rate υ takes on a range of intermediate values bistability may occur: the work in the buffer resides for a while at low values, then switches to a much higher value with the buffer nearly full; after a time there is a switch again to low values, then back to high values, etc. It is shown that exponential backoff can alleviate the unsatisfactory behavior of full buffers and bistability.

Section 3 also contains models for the response time of a tagged packet sent to D in the environment of the other packets contending for space in D's buffer also being retransmitted. It is found that there is a best retransmission interval δ. Too large a retransmission time δ means too much time is wasted in recovery from a lost packet. Too small a retransmission time δ leads to a filling of D's buffer due to retransmission by other sources and results in increased blocking of the tagged packet.

In summary, probabilistic models have been constructed to study the effect of packet size and message size on system throughput and response time of a data-transfer operation. The effect of packet retransmission on response times has also been evaluated.

DEFENSE DATA NETWORK (DDN) PERFORMANCE ANALYSIS USING PROBABILITY MODELING

D. P. Gaver
P. A. Jacobs

1. INTRODUCTION

The DDN is a large packet switching network that services elements of the U.S. DoD. It currently has on the order of 260 packet switching nodes (PSNs) that provide entry to the network. Each PSN is accessed from a number of local hosts, into which traffic flows on the way into, and out of, the DDN itself. The actual users are connected to the hosts via their terminals. A schematic of the relationships appears as Figure 1.

Although the general levels of traffic, in terms of demands from users for capacity to transfer bits from their terminals and back, are known generally, this demand varies randomly in time to a considerable degree. It is apparently not practical to insist that demands for, say, transfer of a large data base from a user at terminal S to another at terminal D be done on a *reservation* basis. Consequently demands occur unpredictably, enter the network as packets, and the packets are routed from PSN to PSN according to an expeditious logic. If too many packets are admitted into the network simultaneously then a detrimental level of congestion occurs at the nodes: buffers that store the incoming datagrams become filled, and the following datagrams, encountering full buffers, are lost and must ultimately be retransmitted. Consequently the transfer of the associated customer's data base is slowed. **Congestion control** measures are required to prevent this from happening, and to compensate for tendencies in the direction of excess congestion, where that congestion comes about in part from the bunchy, random nature of demands for service, i.e., data transfer.

The emphasis of this report is to perform probabilistic analysis of certain features of the DDN system with a view to enhancing or "optimizing" measures of service such as data base throughput and the waiting times experienced by data-base-transfer customers. The particular questions addressed, and the models constructed, are in response to tasking statements supplied by personnel from the Defense Communications, Engineering Center, Reston, Virginia.

2. PACKET SIZING

When a data transfer is to occur in a network like DDN the total collection of bits that comprise the data base (db) is divided into **packets**, i.e., subcollections of contiguous bits from the db plus a header consisting of the bits carrying address information. Packets from a Source terminal are sent into the network *via* a host's PSN, and proceed along individual trajectories, i.e., *via* intermediate PSN's, to the PSN attached ultimately to the Destination's terminal. The trajectories or paths through the network are dictated by a routing algorithm with the general aim of minimizing transit time; individual packets may well follow different paths. The number of PSN's encountered by a packet during its transit from S to D is called the **number of hops**. When a packet reaches its destination it is typically **acknowledged**: information is returned to S to indicate that the particular packet has arrived.

2.1 Packet Size and Error Rate: Model 1

As an individual packet proceeds from S to D it utilizes trunks and passes through buffers and is processed by switches at the PSNs hopped through. Each bit in the packet is susceptible to corruption by various sources along the way. If any bit in a packet is so corrupted the entire packet is viewed as useless and must be retransmitted. If bit corruption is viewed as a process that either does, or does not, occur independently from bit to bit then it is clear that a large packet is more likely to suffer contamination than is a small one. On the other hand each packet of no matter what size has about the same number of header bits, so small packets require the transfer of an uneconomically large amount of header information as compared to the "live" bits that are the data to be transferred. The *question is*: how to choose the number, b, of live bits per packet when a) the success probability per bit is p, and b) the number of header bits per packet is fixed and equal to h?

Note that the above only recognizes one issue that may govern the number of live bits per packet. It focuses on error corruption only; if p is high, as will often be true, the optimal packet size will be large if only errors are important, as may be true when the network is very lightly loaded. Under heavy load conditions a large packet will prove to be non-optimal since large packets tend to be awkward for the PSN's buffers to handle. The latter issue will be addressed later in this report.

Model 1. Mathematical Formulation

A simple analysis of Model 1 is possible. Let

p = probability that a bit is successfully transmitted error-free from S to D. Successive bit conditions are independent; i.e., errors are like biased coin flips (are modeled as Bernoulli trials).

h = number of bits in a header (and tailer); assumed constant.

b = number of "live" bits, a decision variable.

Let N(b) be the (random) number of bits that must be transmitted until a full packet of b live and h header bits arrives intact after one hop. As soon as such a packet arrives at D at the end of the hop another is instantly sent (acknowledgements are assumed to be instantaneous, or else the number of bits is incorporated in h). Clearly N(b) is the inter-event time in a discrete time renewal process; see Ross (1989). The sender receives a **reward** of b bits whenever N(b) terminates, meaning that at independently identically distributed random intervals ($N(b)_j$, j = 1, 2, ...) a packet containing b live, uncontaminated bits arrives. Of course if p is realistically large there will be long strings of good all-live-bit packets, rarely punctuated by very short strings of contaminated packets.

From the following conditional probability argument we can derive what is needed. Note that

$$N(b) = \begin{cases} b+h & \text{with probability } p^{b+h} = e^{-\lambda(b+h)} \\ b+h+N^{\#}(b) & \text{with probability } 1-p^{b+h}=1-e^{-\lambda(b+h)} \end{cases} \qquad (2.1)$$

where $\lambda = -\ln p$.

The first line of (2.1) says that a (new) reward of b bits is collected after one b+ h -bit packet arrives if all bits are uncontaminated. The second line says that a reward of b is collected after the first (contaminated) packet is received *plus* the number of additional packets until an uncontaminated b+h -bit packet arrives; $N^{\#}(b)$ has the same distribution as does N(b) *unconditionally*: after the passage of b+h contaminated packets the process "starts over."

Now take expectations through (2.1):

$$E[N(b)] = (b+h)p^{b+h}+\left\{(b+h)+E[N^{\#}(b)]\right\}(1-p^{b+h}) = (b+h)+E[N(b)](1-p^{b+h}). \quad (2.2)$$

so

$$E[N(b)] = \frac{b+h}{p^{b+h}} \equiv (b+h)e^{\lambda(b+h)}. \quad (2.3)$$

Now the long-run rate of accumulation of good≡live (uncontaminated) packets per packet sent is

$$r(b) = \frac{b}{E[N(b)]} = \frac{b}{(b+h)e^{\lambda(b+h)}}. \quad (2.4)$$

To find the optimal b differentiate r(b) with respect to b and set the derivative equal to zero (despite the fact that b is discrete the result is accurate)

$$\frac{dr(b)}{db} = \frac{[(b+h)e^{\lambda(b+h)}]-b[e^{\lambda(b+h)}+\lambda(b+h)\,e^{\lambda(b+h)}]}{((b+h)e^{\lambda(b+h)})^2}.$$

Cancel $e^{\lambda(b+h)} > 0$ and simplify to obtain the simple quadratic:

$$b^2+hb-h/\lambda = 0,$$

the appropriate solution of which is

$$b_{opt} = \frac{\sqrt{h^2+4h/\lambda}-h}{2}, \quad (2.5)$$

If $p \approx 1$ then λ is small and positive, which sends b_{opt} towards large values:

$$b_{opt} \sim \sqrt{\frac{h}{\lambda}} \tag{2.6}$$

as $\lambda \to 0$.

Model 2.

To generalize the previous model to packets traveling J hops, let $A_i(b)$ be the number of *additional* bits needed to transmit a packet of length b over the i^{th} hop, i = 1, ..., J. We assume that successful transmission of bits over each hop are independent events

$$A_i(b) = \begin{cases} 0 & \text{with probability } p^{b+h} \\ b+h + A_i^{\#}(b) & \text{with probability } 1-p^{b+h} \end{cases} \tag{2.7}$$

where $A_i^{\#}(b)$ is the number of additional bits needed to transmit a packet of length b over the i^{th} hop after the first failed transmission. Letting $\lambda = -\ln p$ as before and taking expectations results in

$$E[A_i(b)] = (b+h)\,(1-e^{-\lambda(b+h)})e^{\lambda(b+h)} = (b+h)\,[e^{\lambda(b+h)}-1]. \tag{2.8}$$

Let $N_J(b)$ be the number of packets needed to transmit a packet of length b over J hops. Then,

$$N_J(b) = b+h+\sum_{i=1}^{J} A_i(b). \tag{2.9}$$

Taking expectations results in

$$E[N_J(b)] = b+h+J(b+h)\,[e^{\lambda(b+h)}-1]. \tag{2.10}$$

The long run average rate at which active bits are accumulated is

$$r_J(b) = \frac{b}{E[N_J(b)]} = \frac{b}{(b+h)[1+J(e^{\lambda(b+h)}-1)]}. \tag{2.11}$$

Taking derivatives with respect to b

$$\frac{d}{db} r_J(b) = \frac{h[1-J(e^{\lambda(b+h)}-1)]-b(b+h)\lambda J e^{\lambda(b+h)}}{(b+h)^2[1+J(e^{\lambda(b+h)}-1)]^2}. \tag{2.12}$$

Setting $\frac{d}{db} r(b) = 0$ results in the equation

$$0 = h[1+J(e^{\lambda(b+h)}-1)]-b(b+h)\lambda J e^{\lambda(b+h)}. \tag{2.13}$$

This equation can be solved numerically to find the maximizing b.

If $\lambda \ll 1$; that is $p \approx 1$; then equation (2.13) can be approximated by

$$\begin{aligned} 0 &= h[1+\lambda J(b+h)]-b(b+h)\lambda J[1+\lambda(b+h)] \qquad (2.14)\\ &\approx h+\lambda Jbh+\lambda Jh^2-b^2\lambda J-bh\lambda J \\ &= h(1+\lambda Jh)-b^2\lambda J. \end{aligned}$$

The maximizing b is

$$b_{opt} \approx \sqrt{\frac{h(1+\lambda Jh)}{\lambda J}} \sim \sqrt{\frac{h}{\lambda J}}. \tag{2.15}$$

As the number of hops increases the maximizing packet size decreases. Note that the above formula is easily generalized to show what happens if λ becomes $\lambda_i = -\ln p_i$, so that the probability of successful transmission depends upon the exact (i^{th}) hop under discussion.

Figures 2 and 3 display the right hand side (RHS) of equation (2.14) for various parameter values. In Figure 2, h = 20 bits and J = 4 with p = 0.99999, 0.99995. The approximate solution (2.15) gives yields $b_{opt} \approx 707$ for p = 0.99999 and $b_{opt} \approx 316$ for p = 0.99995. As expected, a larger probability of successful transmission indicates a larger best packet size. In Figure 3, p = 0.99999, h = 20 and the number of hops is allowed to vary. The approximate formula (2.15) in these cases gives $b_{opt} \approx 1414$ for $J = 1$, $b_{opt} \approx 707$ for $J = 4$, and $b_{opt} \approx 500$ for $J = 8$. Not surprising, the best packet size decreases as the number of hops increases.

Comment. The above analysis tacitly assumes that success probability, p, or p_i on link i, is known; it is also assumed that errors obey Bernoulli-trials laws. Neither is necessarily correct. An adaptive procedure might well be developed to react to changes in p or p_i, particularly to indications of considerable *decrease*, evidence for which could come from demands for retransmit that are not related to losses resulting from encounters with full buffers. A rule that automatically contracts b when an error occurs that compels retransmit could be used. Likewise, a rule that expands b when no errors occur could be devised and evaluated. These problems could be taken up in future work.

3. MODELING A BUFFER THAT EXPERIENCES LOSSES AND RE-TRANSMISSIONS

3.1 Problem Formulation

We consider an aspect of the problem of data transfer in a packet switching store-and-forward network similar to the Defense Data Network (DDN). A source node, S, at which the data resides, wishes to transfer it to destination node D. It must do so via various intermediate nodes. The data at S is packetized, and individual packets are sent forth; if a packet is not acknowledged by a certain time, δ, the packet is re-transmitted, and so on until acknowledgement *is* received. The reason for this tactic is that packets may be lost in transit; loss may be caused by a packet's encountering a full buffer at some intermediate node.

It is intuitively clear, and has been recognized widely, that an overly-short re-transmit time can clog the system with superfluous packets, bringing about drastic slowdowns; see Jacobson (1988), Zhang (1986) and Gerla and Kleinrock (1980). On the other hand, overly-lengthy re-transmit intervals can induce unnecessarily long or prolonged delays. It is of interest to formulate models that suggest a reasonable compromise position, and this is done here. The approach is admittedly approximate, but may form the basis for a better understanding of the situation, and for more realistic strategies. Related issues have been studied for other types of networks; cf. Heyman (1982, 1986) and Fredericks and Riesner (1978).

Our analysis proceeds by first making a very simple fluid model for the contents of the buffer at D. Essentially we are characterizing the *environment* encountered by a packet, P, or *tagged packet*, i.e., a representative packet being sent from S to D as part of the data transfer of interest. Note that we assume that the other packets that contend with P for buffer space *also* re-transmit and occupy the buffer at D. It is their fate to influence, and be influenced by, the magnitude of δ. Thus if δ is large, we expect relatively few superfluous packets at buffer D, and relative ease of entry for P *provided* she is not lost in transit. Decreasing δ tends initially to speed the arrival of a P-copy to D, but if such decrease is universal, applying to all users of the buffer, it also implies that D accumulates more copies of other packets in storage, tending to overload D, leading to rejections of P-copies until one luckily squeezes in. Very likely the time for ultimate squeeze-in will be long if many are in competition. We seek for a reasonable compromise between the too-long δ that delivers excessive delays for lost packets and the too-short δ that leads to an atmosphere of "global warming," losses, and possible system crash. We also model the behavior of *exponential backoff*, a device that increases successive time-out intervals, possibly from δ to 2δ, 2δ to 4δ, *etc*. See Jacobson (1988) and Aldous (1987).

3.2 A simple Fluid Model

(Model 1)

Let w(t) represent the *work* present in D's buffer at time t after the process start; let B be the size of that buffer. By *work* we mean the number of time units required to serve, process, or forward those packets in the buffer at t; B is measured in work units. View w(t) as the expectation of W(t), a random process that would be an unwieldy generalization of the Takacs virtual work process for the M/G/1 queue; see Takacs (1962). In Gaver (1964) a related Takacs model is considered.

We propose this initial model: that if dt>0, and $0 \le w(t) \le B$,

$$w(t+dt) = w(t) + \left[\lambda dt\left(1-(w(t)/B)^p\right)\right]s[1+\upsilon w(t)] - \frac{kw(t)dt}{1+kw(t)}$$

or, letting $dt \to 0$,

$$\frac{dw}{dt} = \lambda s\left(1-(w(t)/B)^p\right)(\upsilon w(t)+1) - \frac{kw(t)}{1+kw(t)}, \quad 0 \le w(t) \le B \qquad (3.1)$$

with $k = \frac{2}{s}$.

Here we think of λdt as being the probability that a new or original packet arrives at the designated buffer, the term $1-(w(t)/B)^p$ being then the probability that it enters D's buffer; we take $p \ge 1$; p can be adjusted to make the probability of loss resemble what is measured. The amount of work the arriving packet brings to D consists of two parts. The first part is the packet's service time s. The second part consists of the amount of work that will arrive due to retransmissions of the arriving packet; there will be $\frac{1}{\delta}w(t) \equiv \upsilon w(t)$ retransmits in the time a new arrival at time t must wait for service, w(t); each retransmission brings in additional work of duration s. We approximate the effect of the total work associated with an arriving packet by adding together the work associated with the original arrival, s, and the total work associated with all retransmissions $\upsilon w(t)s$. Of course work is removed from the buffer at constant, unit, rate when $w(t) > 0$, and at zero rate when $w(t) = 0$. It is convenient to model this discontinuous effect by $kw(t)/(1+kw(t))$, a device apparently originally suggested by C. Agnew (1976); see also Filipiak

(1988); the constant k is chosen so that the limiting solution to (3.1) as $t \to \infty$ when $B = \infty$ and $\upsilon = 0$ is the average work in queue for the M/G/1 queue.

The resulting differential equation, (3.1), for w(t) is non-linear and of first order, and can be solved numerically. Certain qualitative features emerge by examining the right side; put $\rho=\lambda s$ and assume $\rho<1$; viewed as a function of w we have

$$R(w;\upsilon)=\rho\left(1-(w/B)^{p}\right)(\upsilon w+1)-kw/(1+kw). \tag{3.2}$$

For υ small, then $R(w;\upsilon)$ has a unique zero at $0 \le w_1 \le B$.

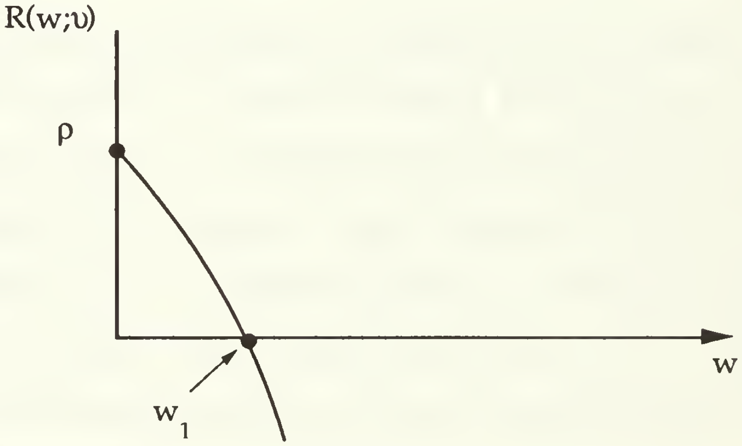


Figure 4

If $\upsilon \gg 1$ (large), then a larger unique zero, w_2, exists.

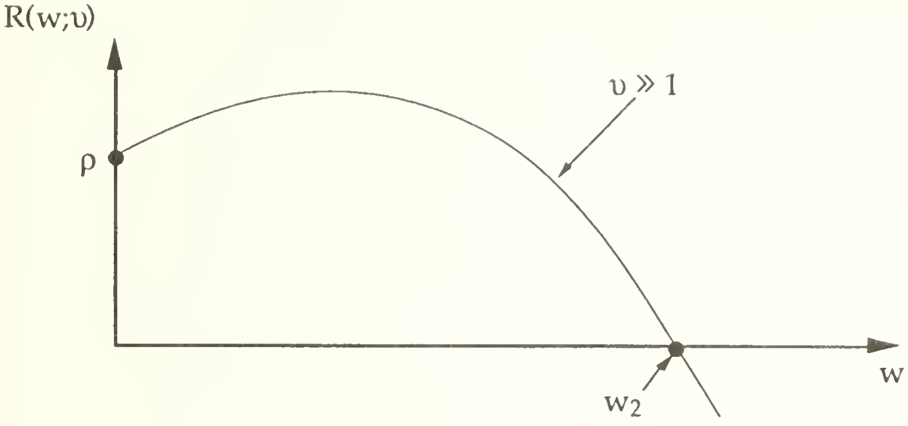


Figure 5

At intermediate values of υ multiple (three) zeros can occur:

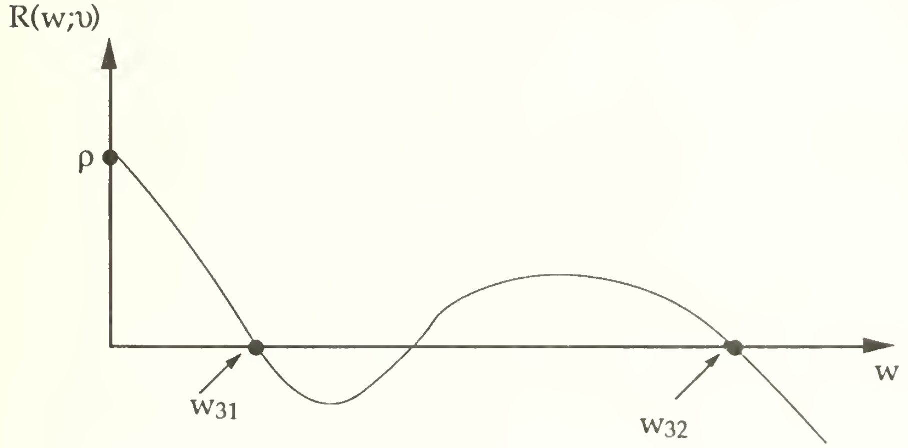


Figure 6

These zeros inform us as to the time-dependent behavior of $w(t)$:

a) If $\upsilon \approx 0$, so re-transmits occur very rarely, then $w(t)$ resembles the mean behavior of an ordinary single-server queue with finite buffer. If $\rho < 1$ (somewhere around 0.5 may be reasonable in practice), then $w(t)$ will tend to be small, i.e., at level w_1.

b) If $\upsilon \gg 1$, so re-transmits occur quite rapidly, then the buffer tends to be quite full, i.e., at level w_2, with w_2 close to B.

c) If υ takes on a range of intermediate values *bistability* may occur: the work in the buffer resides for a time at a low value, w_{31} in the long run, then switches to a much higher value, w_{32}; after a time there is a switch back to (near) w_{31}, then back to w_{32}, etc.

The values w_{31} and w_{32} in Figure 6 are local stable points because if $w(t) = w < w_{31}$ then $R(w;\upsilon) = dw/dt$ is positive, which means that $w(t+dt) > w(t)$, sending $w(t+dt)$ *up* towards w_{31}; if $w(t) = w > w_{31}$ (slightly above) then $R(w;\upsilon) = dw/dt$ is negative, sending $w(t+dt)$ *down* towards w_{31}. Consequently $w(t)$ tends to remain near the stable equilibrium w_{31}. If chance events send the buffer contents sharply higher than w_{31} then the same behavior will occur at w-values near w_{32}, another locally stable equilibrium. The remaining zero, between w_{31} and w_{32}, is seen to be an unstable equilibrium point: any value of w just to its left is associated with a negative derivative, sending system state towards w_{31}; any value just to its right pushes that state toward w_{32}. Bistability has also been noted in models of circuit switched communication networks; Gibbens, et. al. (1988).

3.3 "The" Long-run Solution

Since for fixed parameters υ, B and ρ, $w(t)$ is bounded and displays no absorbing states, it is plausible that there be a long-run solution $w = \lim_{t\to\infty} w(t)$ that satisfies the equation obtained by letting $dw/dt=0$ in (2.1); i.e.,

$$\rho\left(1-\left(\frac{w}{B}\right)^{p}\right)(\upsilon w+1)-\frac{kw}{1+kw}=0. \tag{3.3}$$

It is natural to take $\rho<1$, perhaps $\rho=0.6$ or less, so that the D– buffer is normally rather lightly loaded.

Let

$$f(w)=\left(1-\left(\frac{w}{B}\right)^{p}\right)(\upsilon w+1) \tag{3.4}$$

and

$$g(w)=\frac{1}{\rho}\left(\frac{kw}{1+kw}\right) \tag{3.5}$$

with $k = \frac{2}{s}$.

Note that $f(0) = 1$, $f(B) = 0$, $g(0) = 0$, $g(B) = \frac{1}{\rho}\left(\frac{kB}{1+kB}\right)$. As a function of w, g is strictly increasing. For fixed service time s, as a function of λ, g increases as λ decreases.

Differentiating,

$$\begin{aligned}\frac{d}{dw}f(w)&=\frac{p}{B}\left(\frac{w}{B}\right)^{p-1}[-\upsilon w-1]+\left(1-\left(\frac{w}{B}\right)^{p}\right)\upsilon\\&=\left(\frac{w}{B}\right)^{p-1}\left(-\frac{p}{B}(\upsilon w+1)-\left(\frac{w}{B}\right)\upsilon\right)+\upsilon\\&=\left(\frac{w}{B}\right)^{p-1}\left(-\frac{p}{B}-w\frac{\upsilon}{B}(p+1)\right)+\upsilon\end{aligned} \tag{3.6}$$

For $p = 1$, f is a quadratic function and

$$\frac{d}{dw}f(w)=\left[-\frac{1}{B}-2\frac{w\upsilon}{B}\right]+\upsilon=-\frac{2w\upsilon}{B}+\left(\upsilon-\frac{1}{B}\right). \tag{3.7}$$

Hence if $\upsilon < \frac{1}{B}$, then f is decreasing on [0,B] and there will be 1 zero for the equation $f - g$; the zero will be close to zero. As υ increases, f is an increasing function of υ. For large υ, $f - g$ will also have 1 zero on [0,B], but the zero will be close to B. For intermediate values of υ, $f - g$ will have 3 zeros on [0,B].

Unfortunately there is no clean crisp way of characterizing the shift from one zero to three, so we rely on numerical calculations.

For $p > 1$,

$$\left.\frac{d}{dw} f(w)\right|_{w=0} = \upsilon$$

and $\frac{d}{dw} f(w)$ decreases as w increases on $[0,B]$ for fixed υ and B. Hence, f is a concave function with a unique maximum on $[0,B]$. The equation $f - g$ will have at least 1 zero in $[0,B]$. For small υ and large υ, $f - g$ will have 1 zero. For intermediate υ, $f - g$ can have 3 zeros. Figure 7 shows graphs of f and g for various values of λ and υ with $B = 10$, $s = 1$, and $p = 20$.

The solutions of (3.3) are precisely the long-run values of $w(t)$ discussed earlier: in the long run the work in the buffer is located near w_1 or w_2 when the re-transmit rate υ is either small or large, but hops between w_{31} and w_{32} when υ is of intermediate magnitude, these hops being caused by random events that are not modeled here. A more detailed stochastic analysis is required to fully explicate the situation. But presumably as υ increases, the stationary density of W, a random variable representing long-run buffer contents, will concentrate at a low value near w_1; its density function will gradually exhibit a vestigial mode near w_{32} which grows in size if υ increases. Finally, if υ becomes much larger, the vestigial mode at w_{31} will vanish to be replaced with a major mode at $w_2 \approx B$. It is perhaps even more important to note that such hopping between modes, new preference for the higher mode (essentially full buffer) can equally well be in response to increase in primary traffic rate λ with δ kept fixed.

3.4 On Mean Packet Response Time

Now consider a tagged packet, P, in the environment modeled. Let α be the probability that the tagged packet is lost (encounters a full buffer) on its way to D. Let $\tau(\upsilon) = E[T_{w(\upsilon)}]$, where $T_{w(\upsilon)}$ denotes the response time of the tagged packet in the environment of υ-rate re-transmits.

Then, by a renewal-type backward argument,

$$\tau(\upsilon) = \alpha\left[\frac{1}{\upsilon} + w_I + \tau(\upsilon)\right] + \bar{\alpha}(w(\upsilon)/B)^p\left[\frac{1}{\upsilon} + w_D + \tau(\upsilon)\right] + \bar{\alpha}\left(1 - (w(\upsilon)/B)^p\right)(w_D + w(\upsilon)) \tag{3.8}$$

where w_I is the delay experienced while P percolates through nodes intermediate between S and D, conditional on its being lost before reaching D; w_D is the time to percolate through to the destination node, and $(w(\upsilon)/B)^p$ is the probability that it is finally lost at D. In the first two aforementioned cases P's response time starts over again after loss, while if the packet finally reaches D, after successfully passing through the intermediate nodes, it experiences the expected delay present at D, namely $w(\upsilon)$ in addition to the total percolation time w_D.

Example. Figure 8 displays the expected response time as a function of the retransmission rate υ for a simple case of the previous model. For the examples presented here $w_I = w_D = 0$ and $p = 1$. The amount of work in the system when the tagged packet arrives at the node, $w(\upsilon)$, is taken to be the steady state solution satisfying the equation (3.1). If equation (3.1) has multiple solutions, the largest solution is picked as a conservative choice. Under these assumptions (4.1) reduces to

$$\tau(\upsilon)=\left[1-\bar{\alpha}\left(1-\frac{w(\upsilon)}{B}\right)\right]\left[\frac{1}{\upsilon}+\tau(\upsilon)\right]+\bar{\alpha}\left(1-\frac{w(\upsilon)}{B}\right)w(\upsilon).$$

Solving results in

$$\tau(\upsilon)=\frac{1-\bar{\alpha}\left(1-\frac{w(\upsilon)}{B}\right)}{\bar{\alpha}\left(1-\frac{w(\upsilon)}{B}\right)}\frac{1}{\upsilon}+w(\upsilon). \tag{3.9}$$

Two curves are presented in Figure 8, one for $\alpha = 0.0$ and the other for $\alpha = 0.1$. The other parameters are $\lambda = 0.5$, $s = 1$, and $B = 10$. In both cases, there is a retransmission rate that minimizes expected response time $\tau(\upsilon)$. The minimizing retransmission rate is larger for $\alpha = 0.1$ than for $\alpha = 0.0$ as is to be expected.

3.5 Fluid Model II: Retries

In order to achieve ultimate simplicity the former fluid model omitted explicit consideration of primary *retries:* those original or new packets that are lost immediately because they encounter a full destination buffer. In order to represent their effect it is necessary to introduce a new state variable, r(t), which represents the number of different packets in the retry state at time t; each of these gives rise to repeated work requests until entry is achieved.

The evolution of r(t) is described by the differential equation

$$\frac{dr(t)}{dt}=\lambda\left(w(t)/B\right)^{p}-\upsilon r(t)\left[1-\left(w(t)/B\right)^{p}\right]-\theta r(t). \tag{3.10}$$

The first term on the RHS represents those original packets that encounter a full buffer and hence enter the retry population. The second RHS term represents the rate of retries from that population that enter the buffer. The third term, $-\theta r(t)$, is a control that depletes the retry population as it grows; this corresponds to a timeout on packets that are unable to make the original entry.

The corresponding equation for the actual buffer contents, w(t), is now

$$\frac{dw(t)}{dt} = (\lambda + \upsilon r(t))\left[1-(w(t)/B)^p\right][1+\upsilon w(t)]s - \frac{(2/s)w(t)}{1+(2/s)w(t)}; \qquad (3.11)$$

the addition of $\upsilon r(t)$ to λ in the first term of the RHS represents the arrivals from the retry population.

Note that it has unfortunately been necessary to expand the state space to incorporate the effect of retries, which will quite likely be rare if designs are conservative.

The long-run or steady-state situation is modeled by putting $\frac{dr(t)}{dt} = \frac{dw(t)}{dt} = 0$; one can then solve for $r \equiv r(\infty)$ from (3.10) and substitute into (3.11) to obtain a single equation for $w = w(\infty)$:

$$0 = \lambda s\left[1-(w/B)^p\right][1+\upsilon w]\left[1+\frac{\upsilon(w/B)^p}{\upsilon\left[1-(w/B)^p\right]+\theta}\right] - \frac{(2/s)w}{1+(2/s)w}. \qquad (3.12)$$

Now we can re-define the functions f and g previously introduced and use them to compute the solutions of (3.12), i.e., the stable points:

$$f(w;\upsilon) = \left[1-(w/B)^p\right][1+\upsilon w]\left[\frac{\upsilon+\theta}{\upsilon\left[1-(w/B)^p\right]+\theta}\right] \qquad (3.13)$$

and

$$g(w) = \frac{1}{\lambda s}\left(\frac{(2/s)w}{1+(2/s)w}\right). \qquad (3.14)$$

Notice that $f(0;\upsilon) = 1$, $f(B;\upsilon) = 0$ for all $\theta>0$, a desirable behavior. Also, since the effect of original packet arrivals is entirely concentrated in g(w) it is clear from examination of Figures 9 and 10 that increase in either λ (arrival rate) or s (packet size) will drive g(w) lower for each fixed w, thus increasing the propensity for the higher stable point, that near B, to make its appearance. The parameters of Figure 9 are $s = 1$, $B = 10$, $p = 20$, $\upsilon = 0.3$, $\theta = 0.05$ and $\lambda = 1$, 0.5, 0.3. The parameters of Figure 10 are traffic intensity $\rho = \lambda/s = 1/3$, $B = 10$, $p = 20$, $\upsilon = 0.3$, $\theta = 0.05$ and service times $s = 0.1$, 1, and 1.5. Of course this is the condition that suggests that retransmission rate υ is too high, and could be reduced to accommodate new conditions. Such a general procedure has been suggested in the past and is termed *exponential backoff*. Exponential backoff has been theoretically treated by Aldous (1987) and others for another type of network. It has also been incorporated into an engineering treatment by Jacobson (1988) and others. The next model incorporates exponential backoff into the previous model. Again, the attempt is to parsimoniously trace the impact of a rule followed universally by *all* users of the facility (buffer) under examination. It may be cogently argued that such a universal rule is *not* fair, since it treats all users similarly, regardless of whether they are "large" or "small." If user priorities are equal then it may seem reasonable to accommodate more quickly to the small users than to the large: a lost or rejected (from the buffer) small user is allowed to retransmit sooner than is a large. On the other hand, if many small users are given this privilege then a few large users may be undesirably delayed. We do not attempt to resolve this issue in what follows.

3.6 Fluid Model III: Exponential Backoff

Exponential backoff allows the retransmit interval to become dynamically adaptive to evidence of increased congestion.

Let n(w) be the largest value of n such that

$$\delta[1+\beta+\beta^2+ \dots +\beta^n] < w \tag{3.15}$$

where δ is the minimum time between retransmissions and $\beta>1$ is the backoff factor; n(w) is the number of retransmissions to occur in a time w. An integral approximation to the LHS of (3.15) is

$$\delta\int_0^{n(w)} \beta^x dx = \delta\int_0^{n(w)} \exp\{x \ln\beta\}dx = \frac{\delta}{\ln\beta}\left[\exp\{n(w)\ln\beta\} - 1\right]. \tag{3.16}$$

Thus, an approximation to n(w) can be found by solving the equation

$$w = \frac{\delta}{\ln\beta}\left[\exp\{n(w)\ln\beta\} - 1\right] \tag{3.17}$$

Hence,

$$n(w) \approx c(w) \equiv \frac{1}{\ln\beta}\ln[1 + w\upsilon\ln\beta] \tag{3.18}$$

where $\upsilon = \frac{1}{\delta}$ is the original retransmission rate.

A fluid model for the work in D's buffer at time t, w(t), with backoff and retries satisfies the system of equations

$$\frac{dr(t)}{dt} = \lambda\left(\frac{w(t)}{B}\right)^p - \upsilon r(t)\left[1 - \left(\frac{w(t)}{B}\right)^p\right] - \theta r(t) \tag{3.19}$$

$$\frac{dw(t)}{dt} = (\lambda + \upsilon r(t))\left[1 - \left(\frac{w(t)}{B}\right)^p\right][1 + c(w)]s - \frac{(2/s)w(t)}{1 + (2/s)w(t)} \tag{3.20}$$

where c(w) is defined in (3.18).

As before, the long run or steady-state situation is modeled by putting $\frac{dr(t)}{dt} = \frac{dw(t)}{dt} = 0$; one can then solve for r = r(w) and substitute into (3.20) to obtain a single equation for w=w(∞):

$$0 = \lambda s\left[1-(w/B)^p\right]\left[1+c(w)\right]\left[\frac{\upsilon+\theta}{\upsilon\left[1-(w/B)^p\right]+\theta}\right]-\frac{(2/s)w}{1+(2/s)w}. \qquad (3.21)$$

Let

$$f(w;\upsilon) = \left[1-(w/B)^p\right]\left[1+c(w)\right]\left[\frac{\upsilon+\theta}{\upsilon\left[1-(w/B)^p\right]+\theta}\right] \qquad (3.22)$$

and g(w) be as in (3.14). Notice that f(0;υ) = 1 and f(B;υ) = 0 for all θ>0. The effect of exponential backoff is contained in f. Examination of Figure 11 indicates that increase in the backoff factor β will drive f(w;υ) lower, thus decreasing the propensity for the higher stable point. The parameters of Figure 11 are λ = 1/3, s = 1, p = 20, υ = 0.3, θ = 0.05 and backoff parameter β = 1.01, 1.5 and 3. Hence exponential backoff will provide some protection against a system crash.

Consider a tagged packet P in the environment with retries and backoff. We will model P's expected response time as follows. Let $m_i = E[T_i]$ where T_i is the additional response time needed after i unsuccessful retransmissions. Using the notation of section 3.4, the following recursive relationship can be written

$$m_i = \alpha\left[\beta^i \delta + w_I + m_{i+1}\right]$$
$$+\bar{\alpha}\left(w(\upsilon,\beta)/B\right)^p\left[\beta^i \delta + w_D + m_{i+1}\right]$$
$$+\bar{\alpha}\left[1-\left(w(\upsilon,\beta)/B\right)^p\right]\left[w_D + w(\upsilon,\beta)\right] \tag{3.23}$$

where $w(\upsilon,\beta)$ denotes the expected amount of work present in D's buffer.

Setting $\gamma = (w(\upsilon,\beta)/B)^p$ and

$$d = \alpha w_I + \bar{\alpha}\gamma w_D + \bar{\alpha}\bar{\gamma}\left[w_D + w(\upsilon,\beta)\right]$$
$$= \alpha w_I + \bar{\alpha} w_D + \bar{\alpha}\bar{\gamma} w(\upsilon,\beta), \tag{3.24}$$

substitution yields

$$m_o = d + \delta[\alpha + \bar{\alpha}\gamma] + [1 - \bar{\alpha}\,\bar{\gamma}]m_1$$
$$= d + \delta[\alpha + \bar{\alpha}\gamma] + [1 - \bar{\alpha}\,\bar{\gamma}]\left[d + \beta\delta[\alpha + \bar{\alpha}\gamma] + [1 - \bar{\alpha}\,\bar{\gamma}]m_2\right]$$
$$\vdots$$
$$= d\left[1 + [1 - \bar{\alpha}\,\bar{\gamma}] + [1 - \bar{\alpha}\,\bar{\gamma}]^2 + \cdots + [1 - \bar{\alpha}\,\bar{\gamma}]^n\right]$$
$$+\delta[\alpha + \bar{\alpha}\gamma]\left\{1 + [1 - \bar{\alpha}\bar{\gamma}]\beta + \cdots + [1 - \bar{\alpha}\,\bar{\gamma}]^n \beta^n\right\}$$
$$+[1 - \bar{\alpha}\bar{\gamma}]^{n+1} m_{n+1}. \tag{3.25}$$

Thus, if $\beta[1 - \bar{\alpha}\bar{\gamma}] < 1$, then the expected response time

$$m_0 = \frac{d}{\bar{\alpha}\bar{\gamma}} + \frac{\delta[\alpha + \bar{\alpha}\gamma]}{1 - \beta[1 - \bar{\alpha}\bar{\gamma}]}. \tag{3.26}$$

If $\beta[1 - \bar{\alpha}\bar{\gamma}] > 1$, then the expected response time is infinite.

Putting $w_I = w_D = 0$, it follows that $d = \bar{\alpha}\bar{\gamma}\omega(\upsilon,\beta)$ and

$$m_0 = w(\upsilon,\beta) + \frac{\delta[\alpha + \bar{\alpha}\gamma]}{1 - \beta[1 - \bar{\alpha}\bar{\gamma}]}. \tag{3.27}$$

3.7 A Diffusion Model

In this section we replace the simple deterministic fluid model of Section 3.2 with a simple stochastic diffusion model. Let $W(t)$ represent the random amount of work present in D's buffer at time t. We will model $\{W(t);\ t \geq 0\}$ with a diffusion process satisfying the following stochastic differential equation

$$dW(t) = a(W(t)) + \sqrt{b(W(t))}\, dU(t) \tag{3.28}$$

where

$$a(w) = \lambda(1-(w/B)^p)\ [1+\upsilon w]s-1; \tag{3.29}$$

$$b(w) = \lambda(1-(w/B)^p)\ [1+\upsilon w]^2 s^2; \tag{3.30}$$

and $\{U(t);\ t \geq 0\}$ is a standard Brownian motion. The infinitesimal mean of $W(t)$, $a(w)$, consists of two parts. When there are w units of work in the buffer a new packet arrives and enters the buffer at a rate $\lambda(1-(w/B)^p)$; the packet brings a total amount of work to the buffer equal to his service time s plus the additional work due to retransmissions which we take to be υws as before. Finally, the server works at a unit rate to clear work from the buffer. The infinitesimal variance of $W(t)$, $b(w)$, corresponds to the variance term for a compound Poisson process with arrival rate $\lambda(1-(w/B)^p)$ and second moment of the magnitude of arriving work which we take to be $(1+\upsilon w)^2 s^2$.

Figures 12-15 present sample functions of the following simulation of a discrete time approximation to $\{W(t);\ t \geq 0\}$. In the following $\Delta = 0.1$ is the time increment and $\{U_n\}$ are independent identically distributed standard normal random variables. The discrete time process evolves as

$$W^*((n+1)\Delta) = W(n\Delta) + a(W(n\Delta))\Delta + \sqrt{b(W(n\Delta))\Delta}\; U_{n+1}$$

where a(w) and b(w) are given by equations (3.24)–(3.25). Finally, the amount of work in the buffer at time $(n+1)\Delta$, $W((n+1)\Delta)$, is $W^*((n+1)\Delta)$ truncated to be within the interval [0,B]; that is,

$$W((n+1)\Delta) = \max(\min(W^*((n+1)\Delta),B),0).$$

In Figures 12–15, the arrival rate of new packets is $\lambda = 0.5$, the service time s = 1, the buffer size B = 10 and p = 20. The initial work in the buffer is W(0) = 1. In Figure 12, $\upsilon = 0.0$ and $\{W(t); t \geq 0\}$ remains less than 2 most of the time with occasional movement to values larger than 3 but less than 4. In Figure 13, $\upsilon = 0.1$ and the behavior of $\{W(t)\}$ is similar to that in the case when $\upsilon = 0.0$. However, the values of $\{W(t)\}$ are somewhat larger. In Figure 14 $\upsilon = 0.5$ and $\{W(t); t \geq 0\}$ is oscillating between small values and large values close to the buffer size of 10. This sample path behavior is called bistability. The behavior was suggested by the multiple zeros of the steady state fluid model equation (3.3). In Figure 15, $\upsilon = 0.8$ and $\{W(t); t \geq 0\}$ is also oscillating between large and small values with more of its time being spent at the full buffer level 10. If the time interval Δ is decreased, similar behavior results but it is less pronounced.

A simulation was done of the work in the system after a packet arrival for Poisson arrivals of the packets. In particular, let S_n be the time of the n^{th} arrival and W_n be the amount of work in the system just after the arrival of the n^{th} customer. Let

$$W^0_{n+1} = \max(\{W_n-(S_{n+1}-S_n)\},0)$$

be the amount of work in the system just before the $(n+1)^{st}$ arrival. The amount of work in the system just after the (n+1) arrival is

$$W_{n+1} = \min(\{W^0_{n+1} + \upsilon(W^0_{n+1} + 1)s\}, B).$$

In the simulations $W_0 = 1$. Figures 16-18 present sample paths of $\{W_n\}$ for $\lambda = 0.5$, $s = 1$, $B = 10$ and various values of υ. In Figure 16, $\upsilon = 0.3$ and the amount of work in the system is small. Figure 17 with $\upsilon = 0.5$ shows some oscillation between small and large values of work. Finally, Figure 18 with $\upsilon = 0.8$ shows a great deal of oscillation with more time being spent with a full buffer. Thus the diffusion approximation is representing the bistable behavior of the system with Poisson arrivals.

3.8 The Density Function for the Steady State Amount of Work in the Buffer

We assume the conditional density function of W(t), given $W(0) = w_0$, satisfies the Fokker-Planck equation

$$\frac{\partial}{\partial t} f(t,w) = \frac{-\partial}{\partial w}[a(w)f(t,w)] + \frac{1}{2}\frac{\partial^2}{\partial w^2}[b(w)f(t,w)] \tag{3.31}$$

where we have suppressed the dependence on w_0.

Integrate once on w, to obtain

$$\frac{\partial}{\partial t} F(t,w) = -a(w)f(t,w) + \frac{1}{2}\frac{\partial}{\partial w}(b(w)f(t,w)). \tag{3.32}$$

Assume a limiting density function exists and as $t \to \infty$ $\lim_{t\to\infty} \frac{\partial}{\partial w} F(t,w) = 0$. Set $f(w) = \lim_{t\to\infty} f(t,w)$. The density function f satisfies the equation

$$0 = -a(w)f(w) + \frac{1}{2}\frac{d}{dw}(b(w)f(w)). \tag{3.33}$$

Solving (3.33) results in

$$
\begin{aligned}
f(w) &= C\frac{1}{b(w)}\exp\left\{\int_0^w \frac{2a(y)}{b(y)}dy\right\} \\
&= C\frac{1}{b(w)}\exp\left\{\int_0^w 2\left[\frac{1}{(1+\upsilon y)s} - \frac{1}{b(y)}\right]dy\right\} \qquad (3.34) \\
&= C\frac{1}{b(w)}[1+\upsilon y]^{2/\upsilon s}\exp\left\{-2\int_0^w \left[\frac{1}{b(y)}\right]dy\right\}
\end{aligned}
$$

where C is a normalizing constant.

Example. Figure 19 shows the density function (3.34) for the case p = 1, $\lambda = 0.5$, s = 1, B = 10 and two values of $\upsilon = 0.1$ and 0.8. The integral of (3.34) is evaluated using partial fractions. For $\upsilon = 0.1$, the mass of the density function is concentrated at small values of w. For $\upsilon = 0.8$, the mass of the density function is bathtub shaped with mass concentrated both at small values of w near zero and at large values of w near the buffer limit of 10. This corresponds to the fluid approximation with two stable points, w_{31} and w_{32}.

4. CONCLUSIONS

In response to the original tasking on this project, mathematical analysis has been conducted that exhibits the influence of packet size on system throughput and waiting time for elements of a data-transfer operation. Congestion control protocols are evaluated that tend to minimize the unwanted effects of retransmits on overall system performance, as measured by response times to request for data transfer.

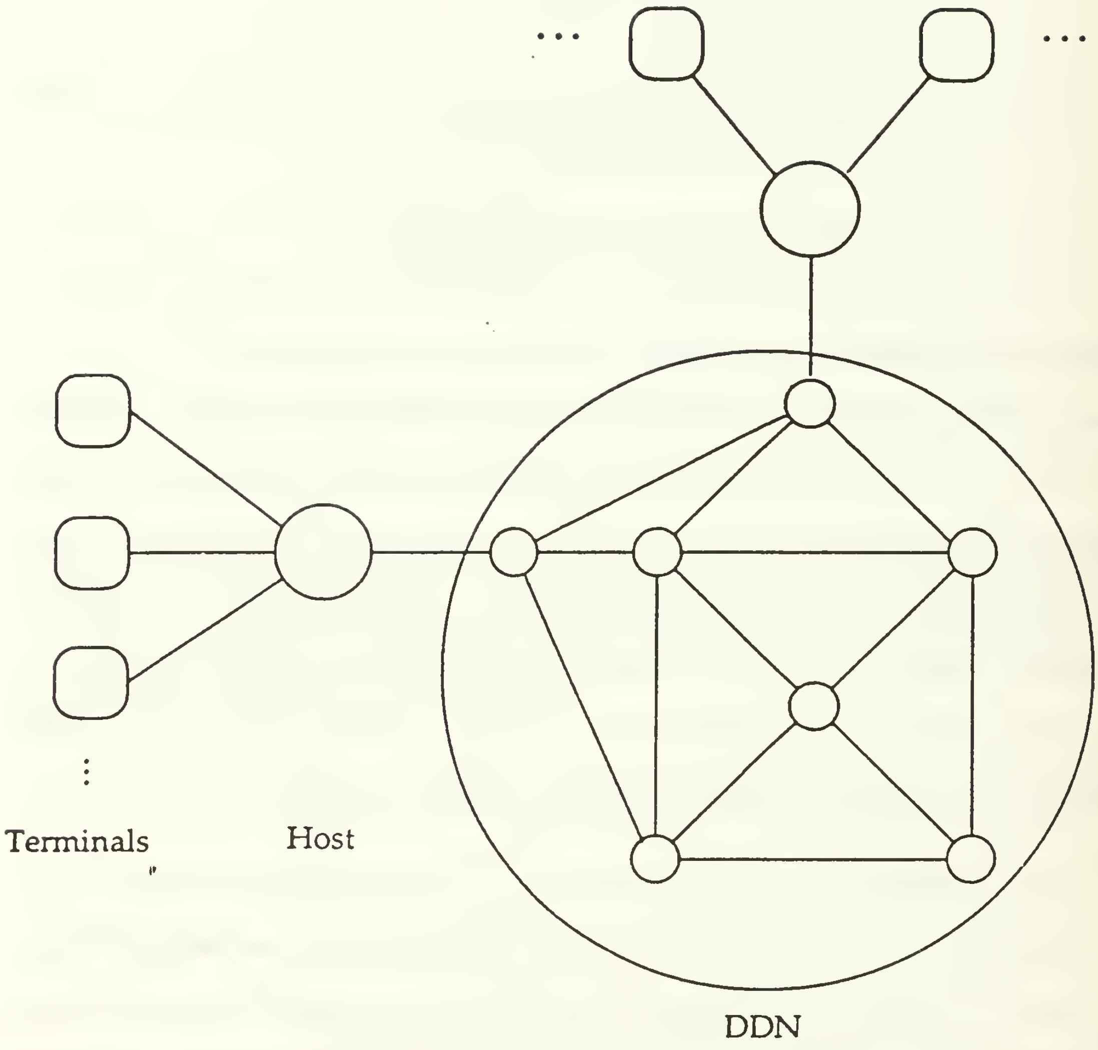


Figure 1

DERIVATIVE OF RATE OF ACCUMULATING ACTIVE BITS

P(SUCCESSFUL BIT TRANS)=.99999,.99995;OVERHEAD=20;HOPS=4

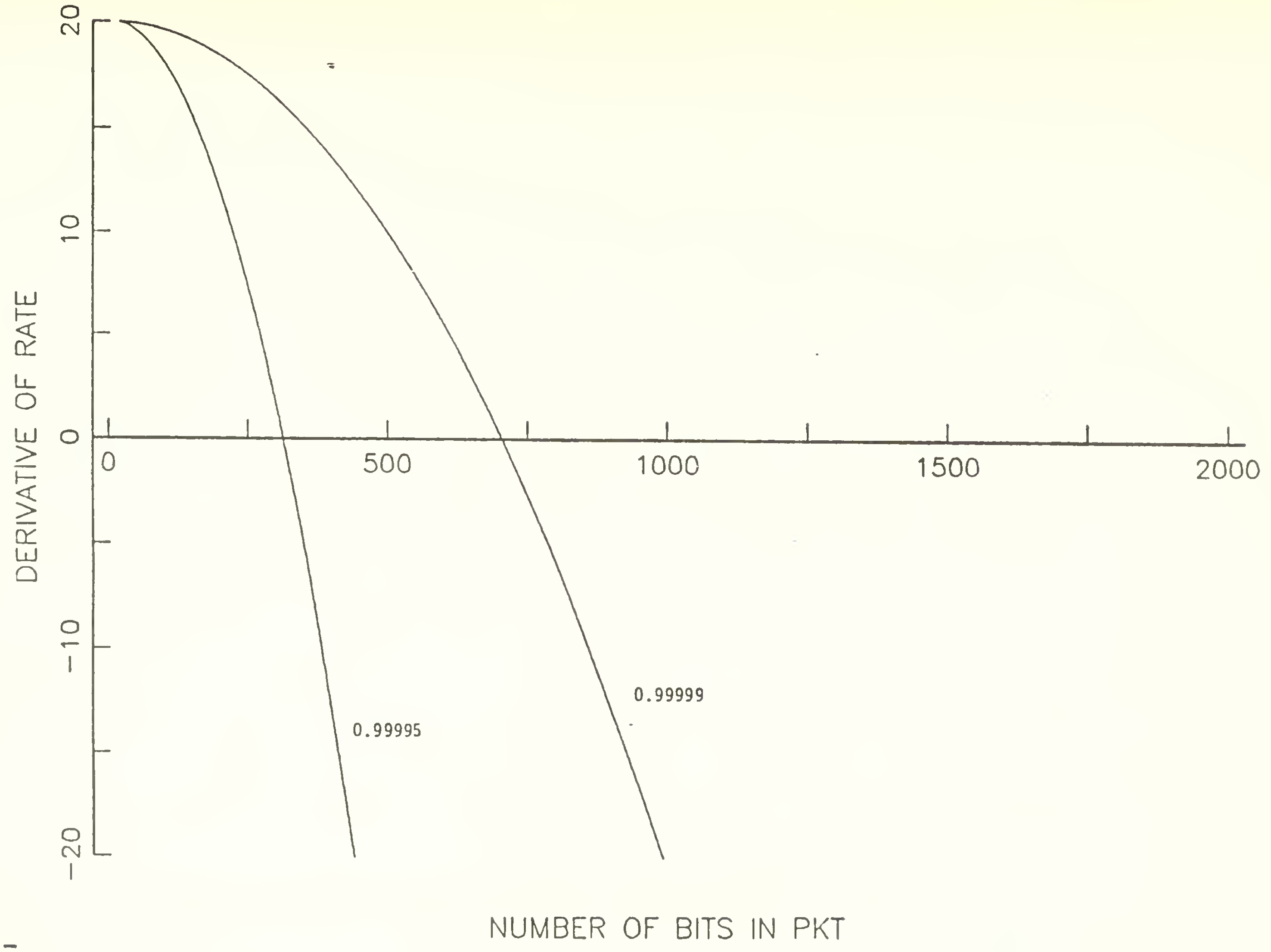


Figure 2

DERIVATIVE OF RATE OF ACCUMULATING ACTIVE BITS

P(SUCCESSFUL BIT TRANS)=.99999;OVERHEAD=20;HOPS=1,4,8

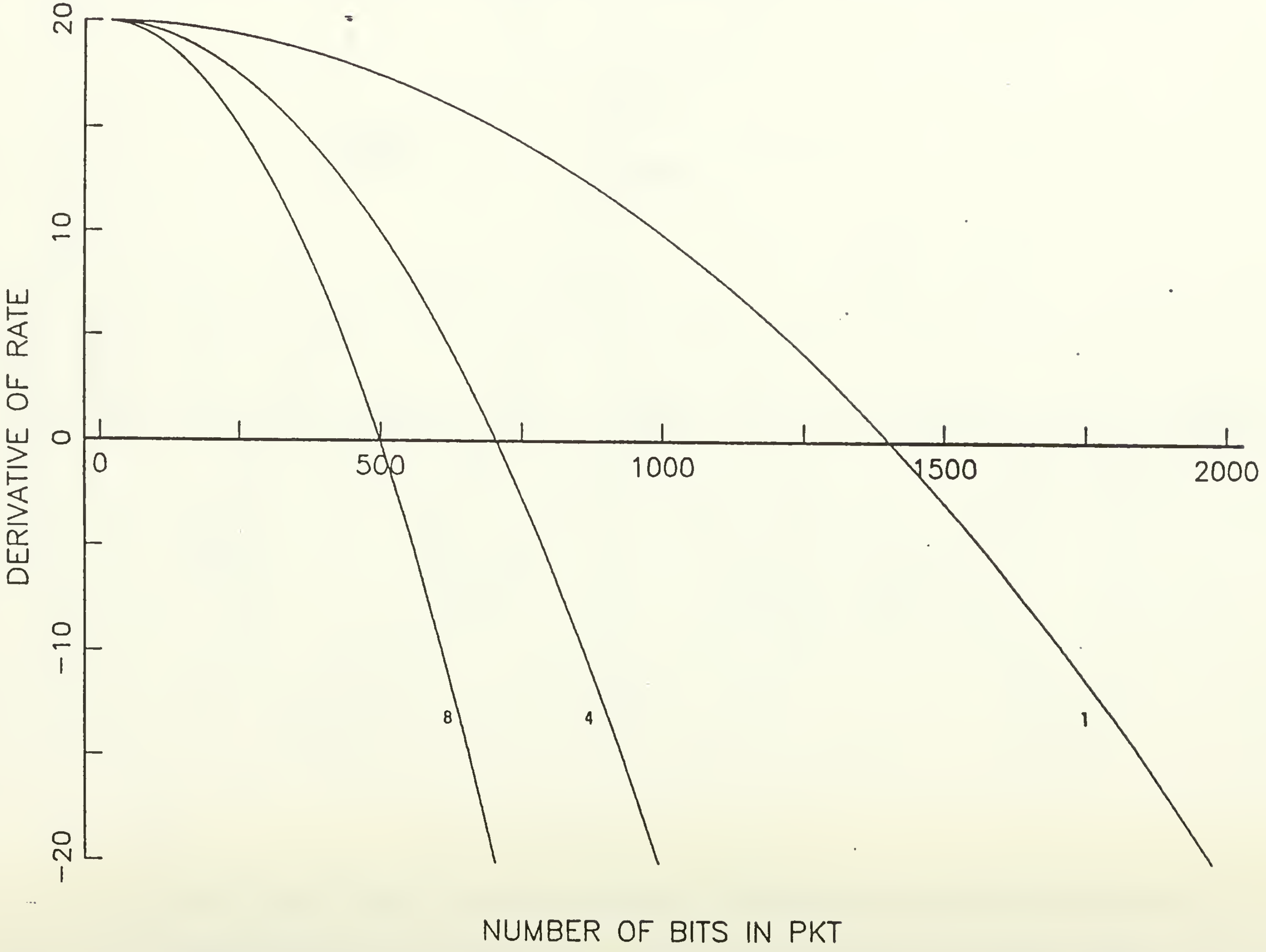


Figure 3

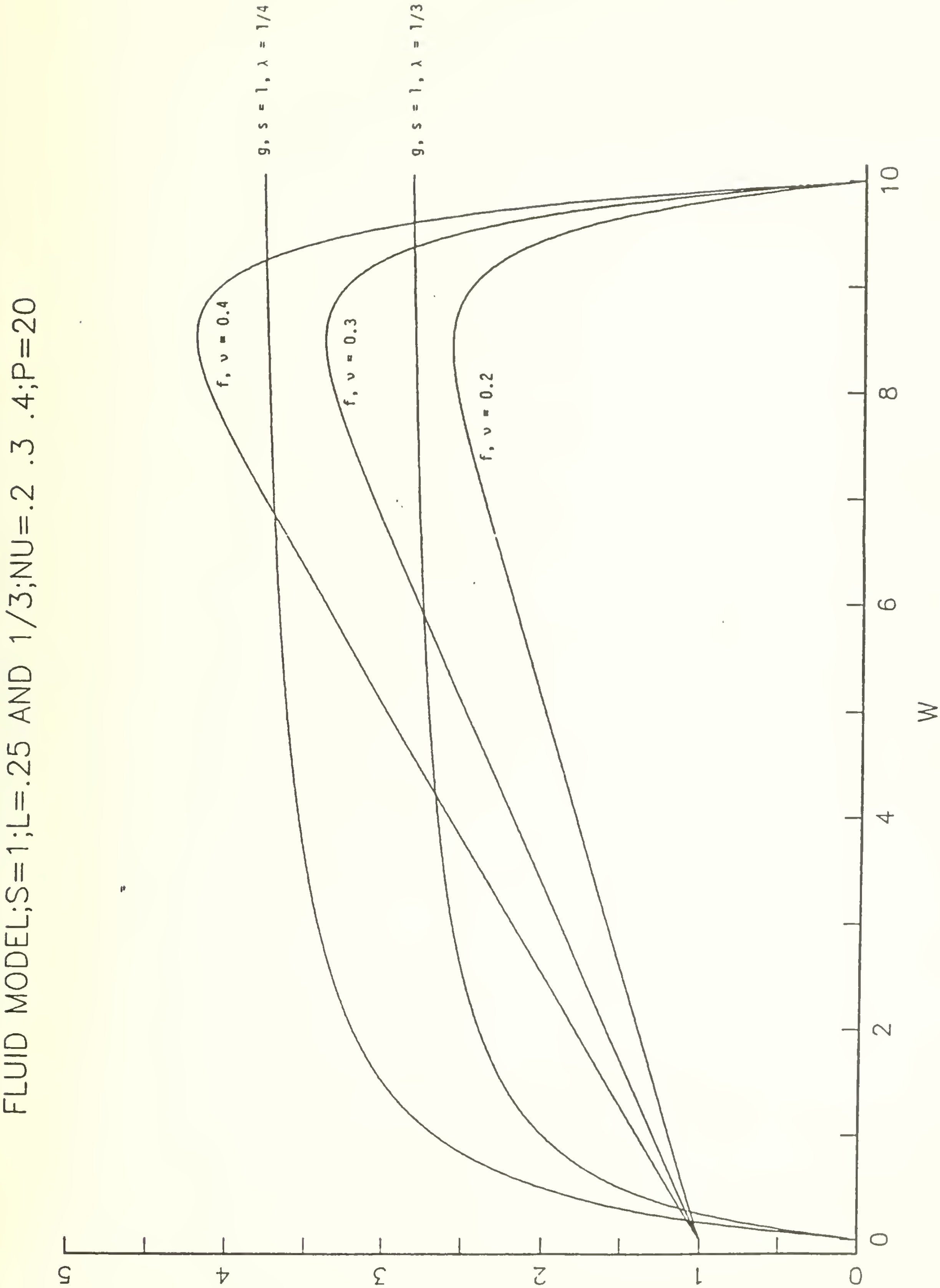

FLUID MODEL;S=1;L=.25 AND 1/3;NU=.2 .3 .4;P=20
g, s = 1, λ = 1/4
g, s = 1, λ = 1/3
f, ν = 0.4
f, ν = 0.3
f, ν = 0.2
W
0
2
4
6
8
10
0
1
2
3
4
5

Figure 7

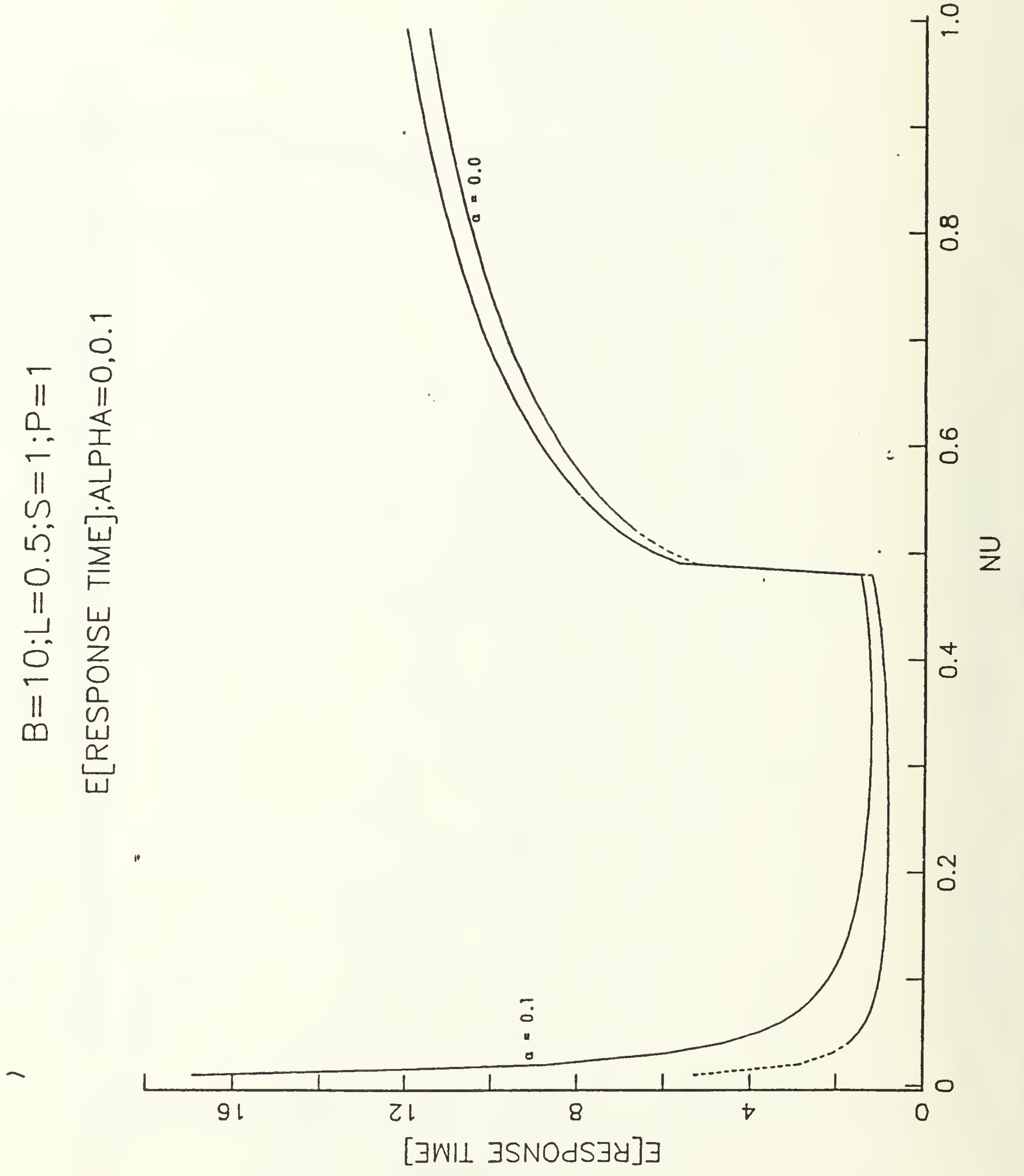

B=10;L=0.5;S=1;P=1
E[RESPONSE TIME];ALPHA=0,0.1
E[RESPONSE TIME]
0
4
8
12
16
NU
0
0.2
0.4
0.6
0.8
1.0
α = 0.1
α = 0.0

Figure 8

FLUID MODEL WITH RETRIES

S=1;P=20;NU=0.3;θ=.05;L=1,0.5,0.3

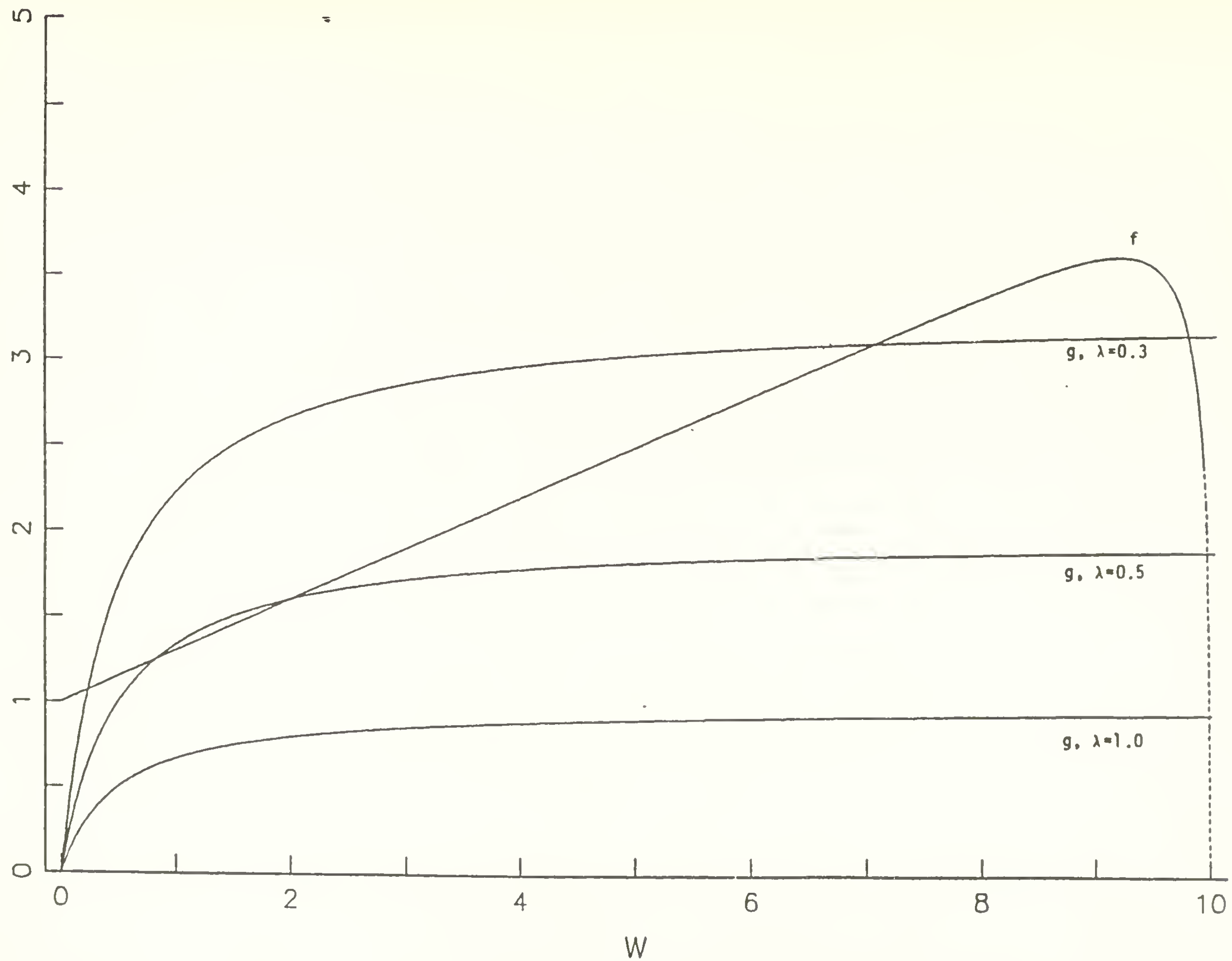


Figure 9

FLUID MODEL WITH RETRIES

$\rho=1/3$;P=20;NU=0.3;θ=0.05;S=0.1,1,5

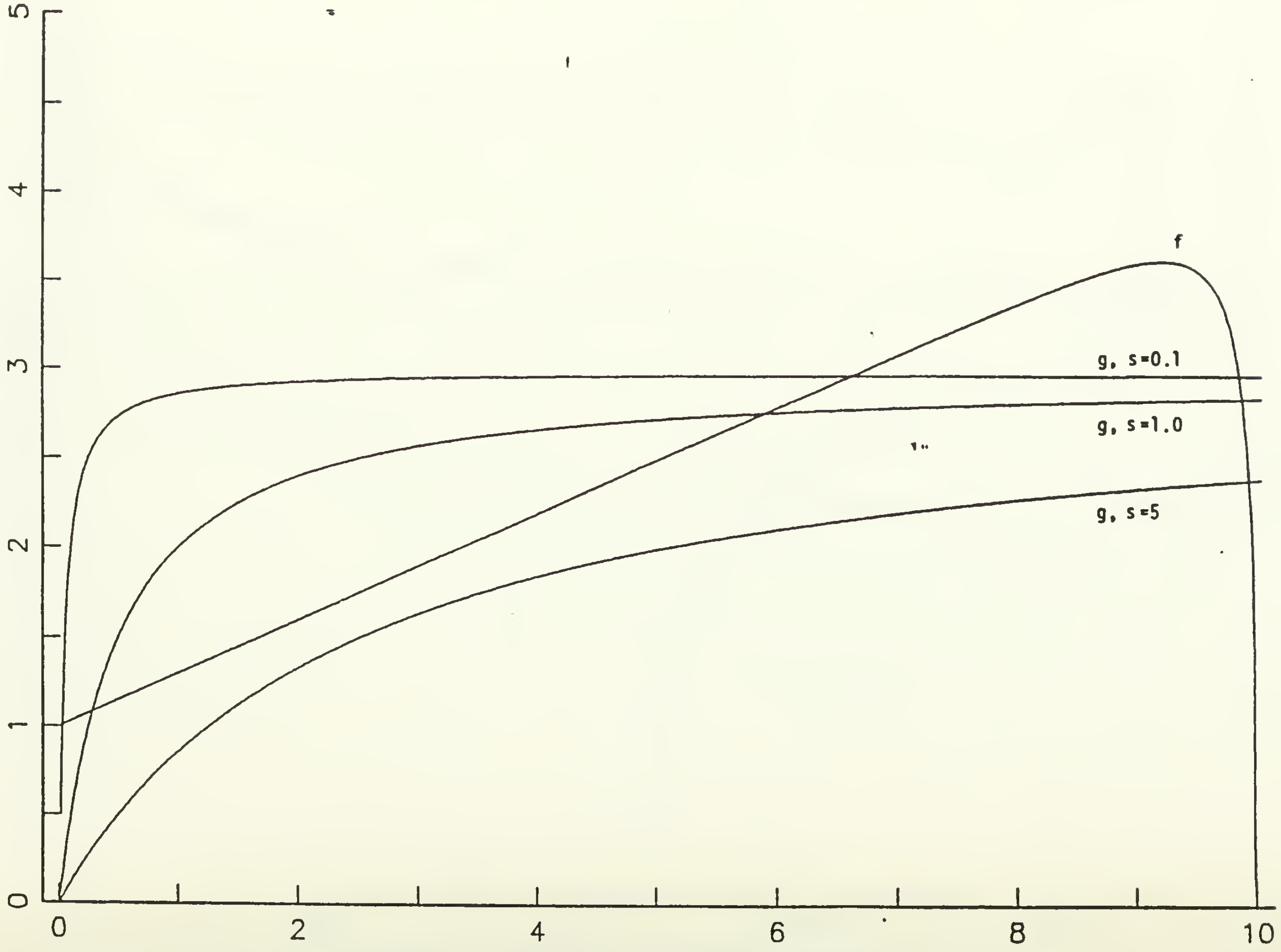


Figure 10

FLUID MODEL WITH RETRIES AND BACKOFF

$L=1/3;S=1;P=20;NU=0.3;\theta=0.05;BETA=1.01,1.5,3$

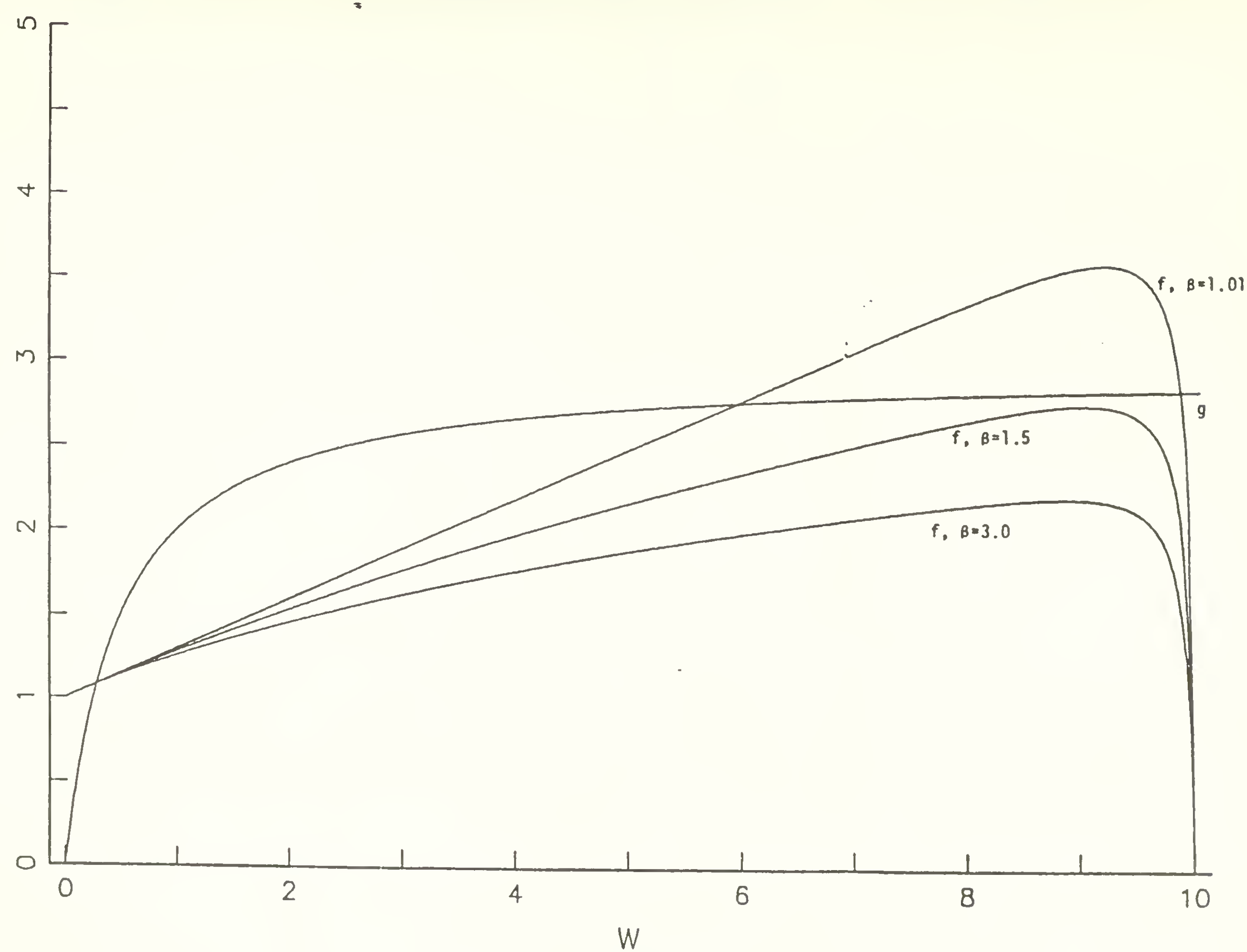


Figure 11

SIMULATION OF DIFFUSION;L=.5;S=1;P=20;NU=0.0
0
1
2
3
0
500
1000
1500
2000
2500
T×10

Figure 12

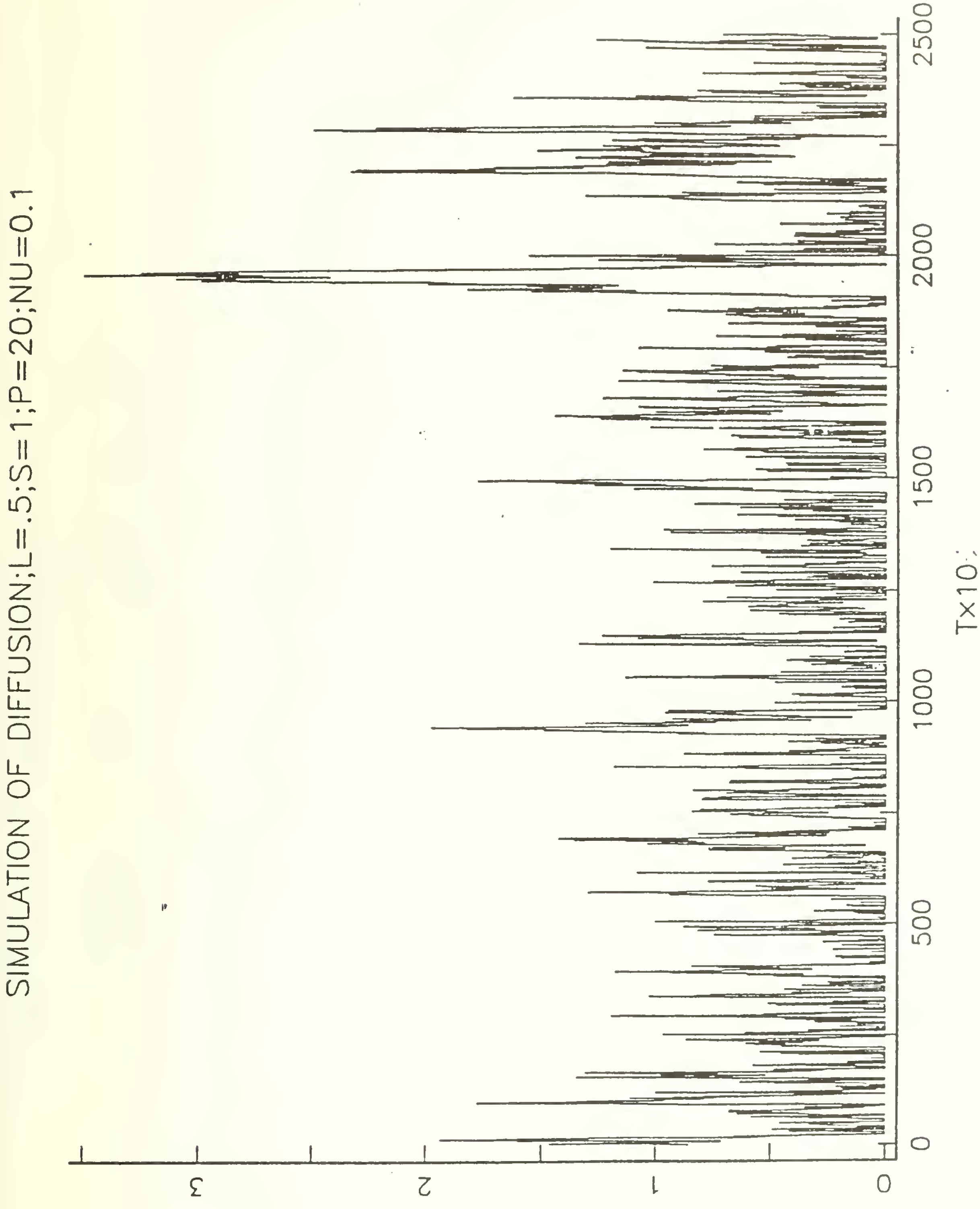

SIMULATION OF DIFFUSION;L=.5;S=1;P=20;NU=0.1
0
1
2
3
0
500
1000
1500
2000
2500
T×10

Figure 13

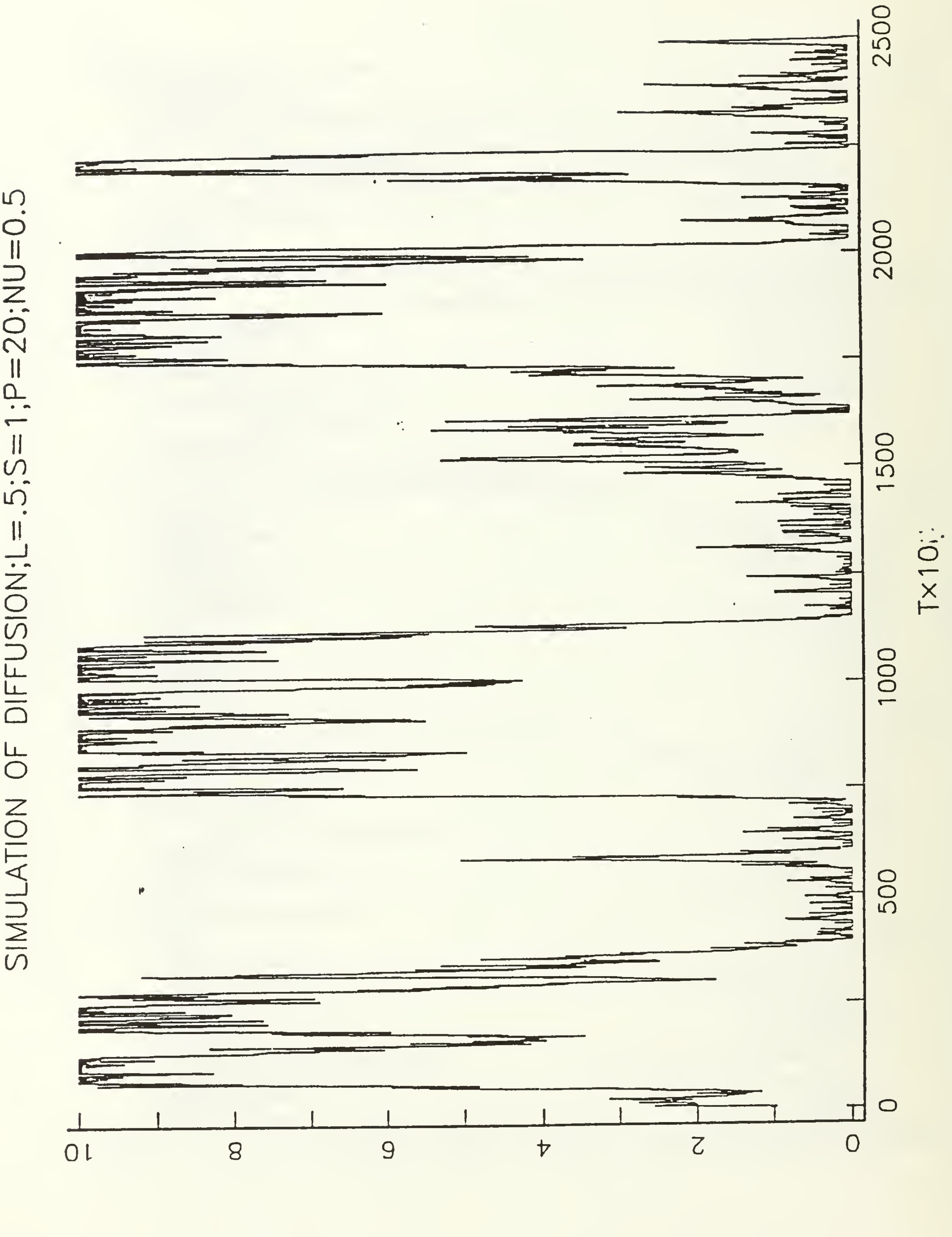

SIMULATION OF DIFFUSION;L=.5;S=1;P=20;NU=0.5
0
2
4
6
8
10
0
500
1000
1500
2000
2500
T×10

Figure 14

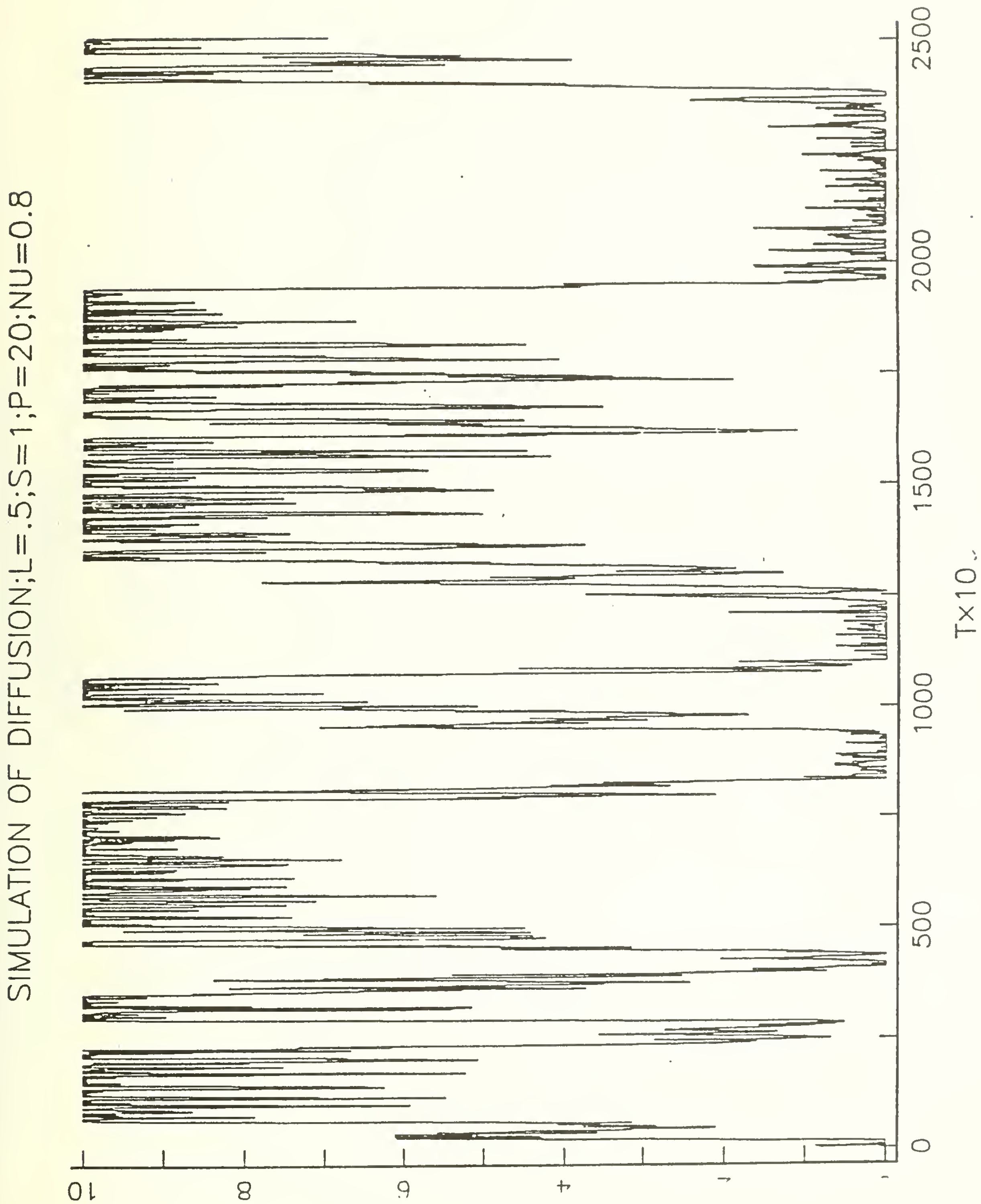

SIMULATION OF DIFFUSION;L=.5;S=1;P=20;NU=0.8
10
8
6
0
500
1000
1500
2000
2500
T×10

Figure 15

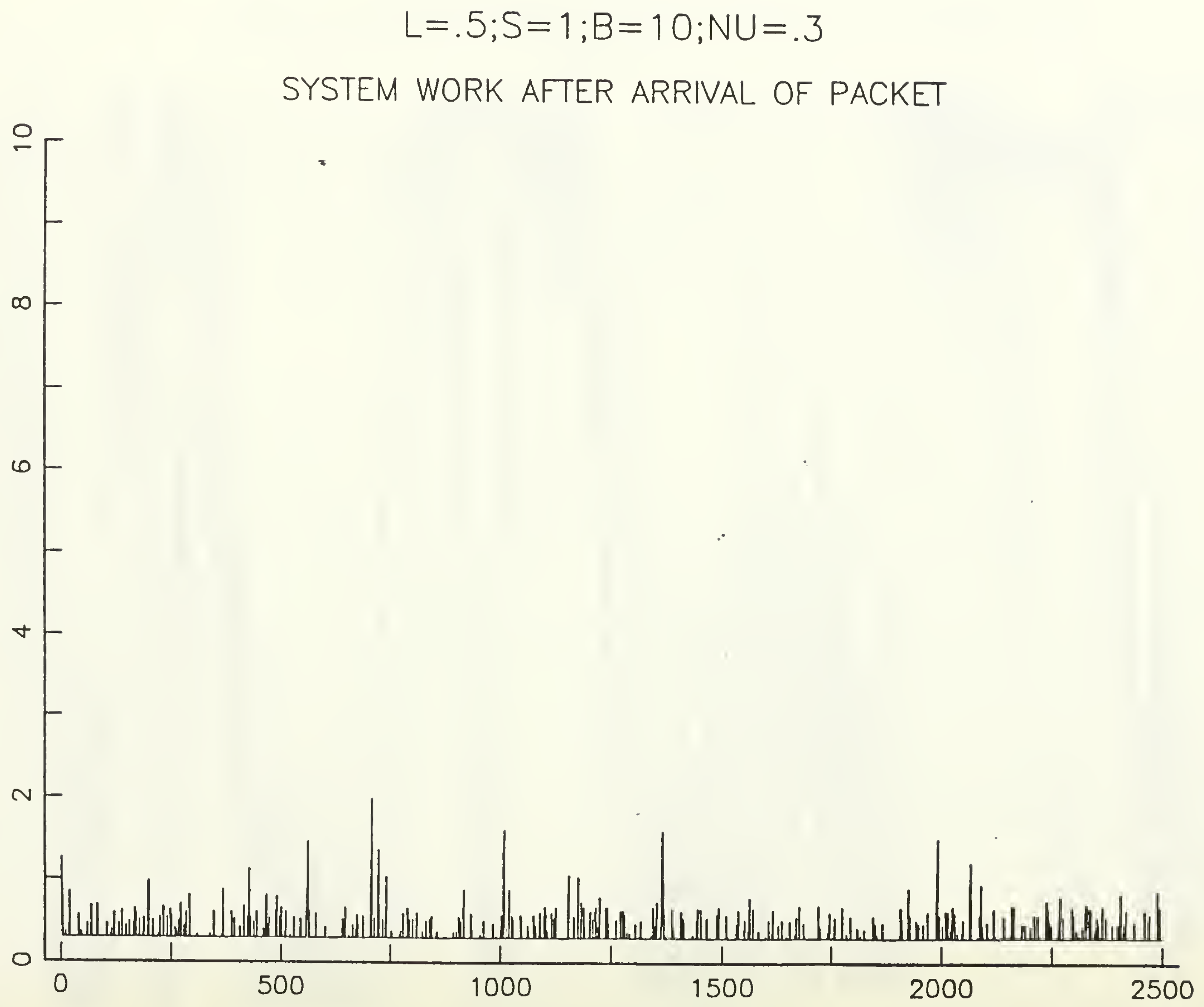
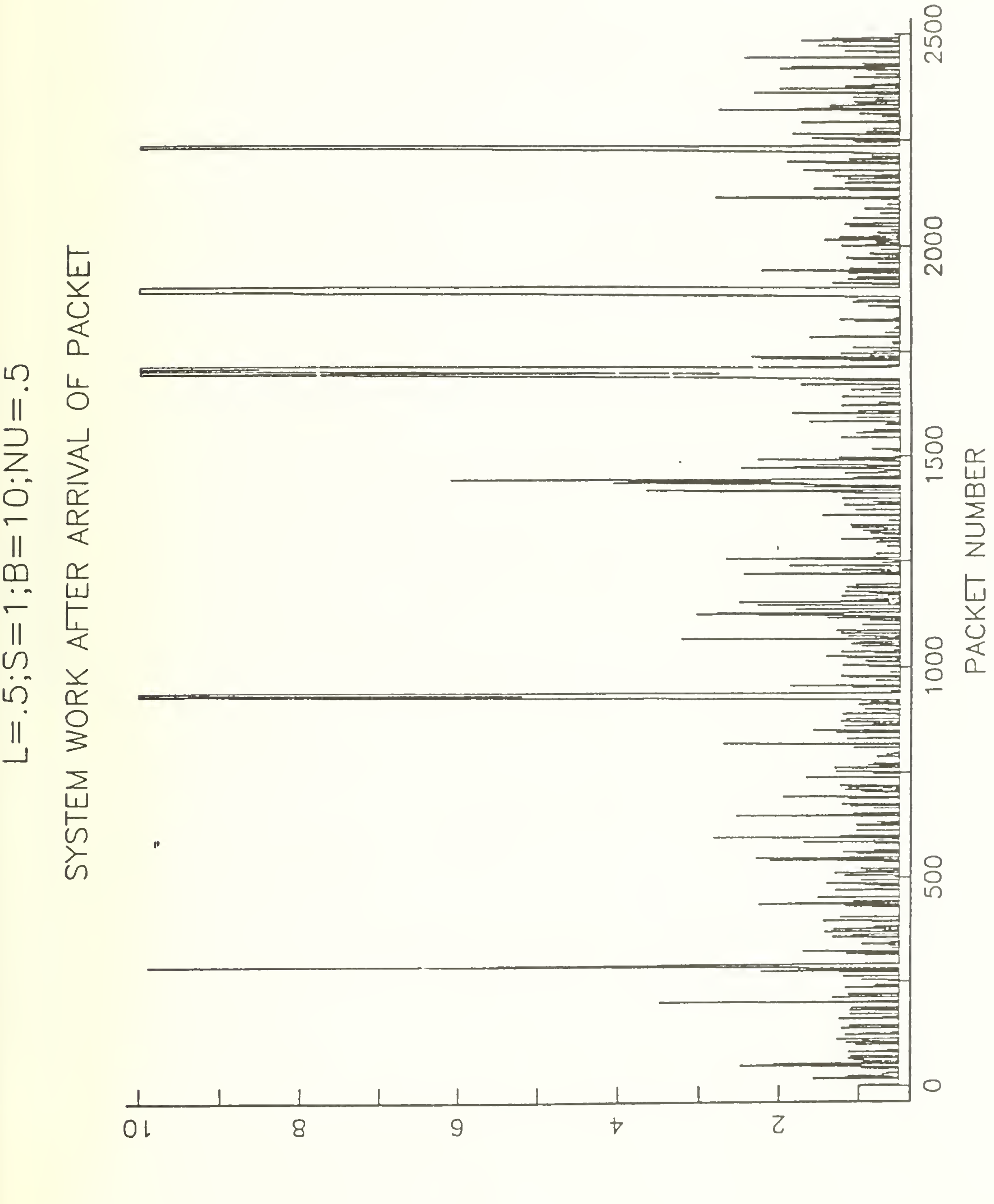
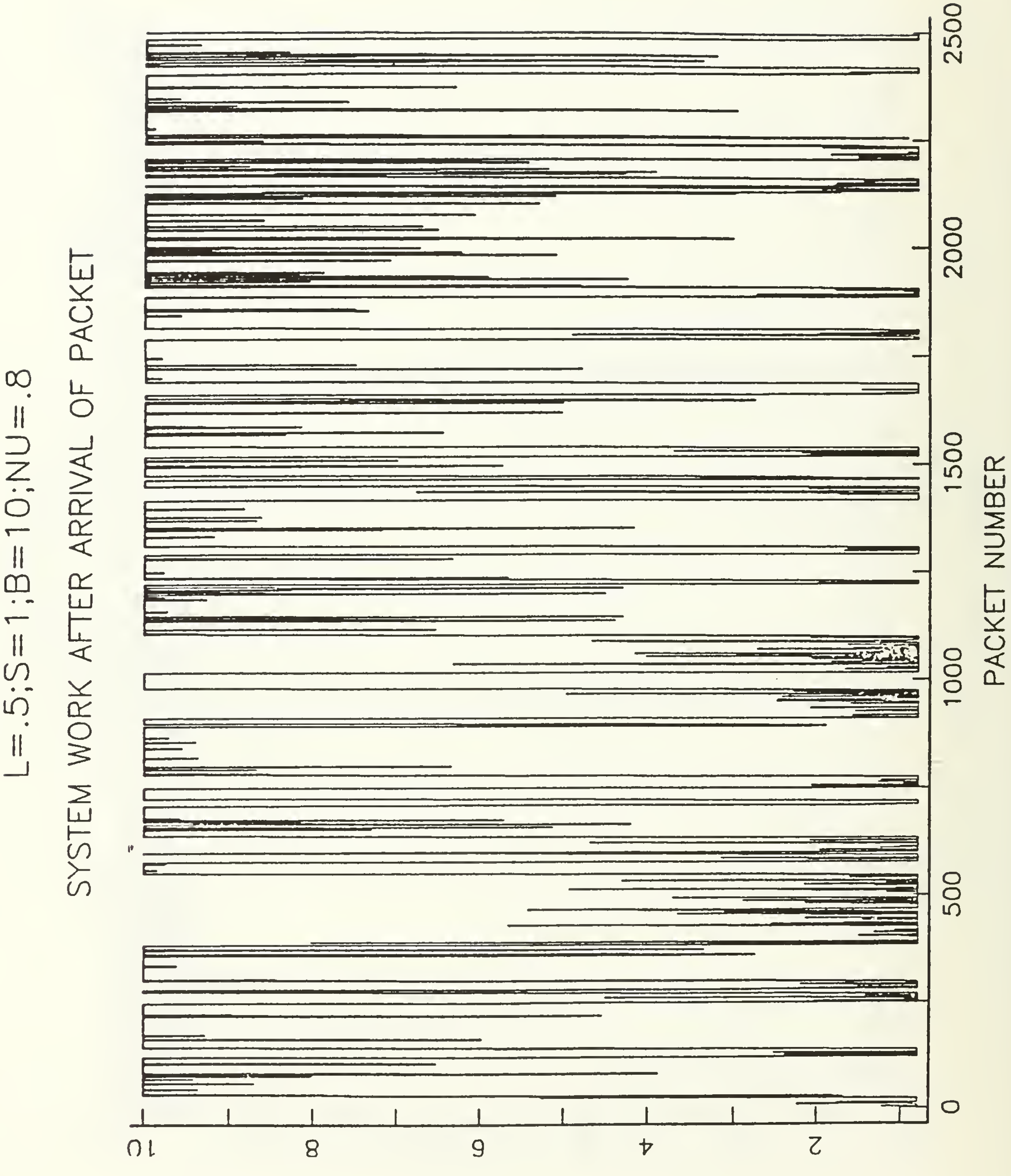

L=.5;S=1;B=10;NU=.3
SYSTEM WORK AFTER ARRIVAL OF PACKET
0
2
4
6
8
10
0
500
1000
1500
2000
2500
PACKET NUMBER

Figure 16



L=.5;S=1;B=10;NU=.5
SYSTEM WORK AFTER ARRIVAL OF PACKET
10
8
6
4
2
0
500
1000
1500
2000
2500
PACKET NUMBER

Figure 17



L=.5;S=1;B=10;NU=.8
SYSTEM WORK AFTER ARRIVAL OF PACKET
PACKET NUMBER
0
500
1000
1500
2000
2500
2
4
6
8
10

Figure 18

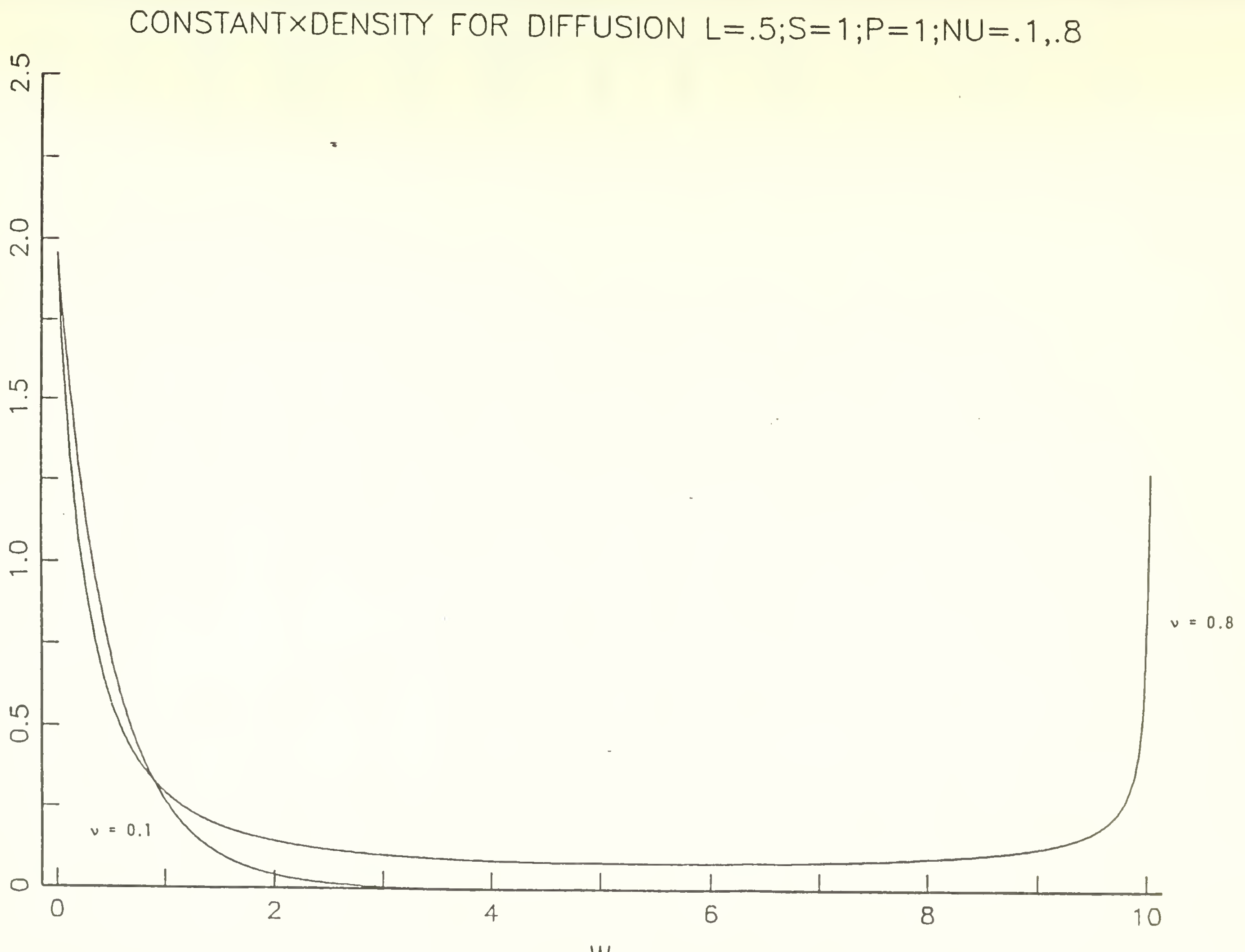

CONSTANT×DENSITY FOR DIFFUSION L=.5;S=1;P=1;NU=.1,.8
2.5
2.0
1.5
1.0
0.5
0
ν = 0.1
ν = 0.8
0
2
4
6
8
10
W

Figure 19

REFERENCES

Agnew, C. W. (1976). "Dynamic modeling and control of congestion prone systems," *Operations Research*, pp. 400-419.

Aldous, D. J. (1987). "Ultimate instability of exponential backoff protocol for acknowledgment based transmission control of random access communication channels." *IEEE Transactions on Information Theory*, IT-33(2), March, pp. 219-223.

Filipiak, J. (1988). *Modelling and Control of Dynamic Flows in Communication Networks*. Springer-Verlag.

Fredericks, A. A. and Reisner, G. A. (1979). "Approximations to stochastic service systems, with an application to a retrial model. *The Bell System Technical Journal* **58** (3) pp. 557-576.

Gaver, D. P. (1964). "An absorption probability problem." *J. Math Analysis and Appl.* **9**, pp. 384-393.

Gerla, M. and Kleinrock, L. (1980). "Flow control: A comparative survey," *IEEE Trans. on Communications*. Vol. COM-28, No. 4.

Gibbens, R. J., Hunt, P. J. and Kelley, F. P. (1988). *Bistability in Communication Networks*. Technical Report, Stochastic Networks Group, University of Cambridge, Cambridge, England.

Heyman, D. P. (1982). "An analysis of the carrier-sense multiple-access protocol." *The Bell System Technical Journal*, **61** (8) pp. 2023-2051.

Heyman, D. P. (1986). "The effects of random message sizes on the performance of the CSMA/CD protocol." *IEEE Transactions on Communications* COM-34 (6) pp. 547-553.

Jacobson, V. (1988). "Congestion avoidance and Control." *Proceeding of SIGCOMM '88*. ACM.

Ross, S. M. (1989) *Introduction to Probability Models*. Fourth Edition, Academic Press, New York.

Takács, L. (1962). *Introduction to the Theory of Queues*. Oxford University Press, New York.

Zhang, L. (1986). "Why TCP timers don't work well," *Proceeding of SIGCOMM '86 ACM August 1986*, pp. 397-405.

APPENDIX A. PROBABILISTIC EVALUATION OF A PACKET SWITCHING PROTOCOL

D. P. Gaver
P. A. Jacobs
P. Purdue

Abstract

A simple initial model is introduced to study the effect of the number of packets in a message upon data transfer time in the Defense Communications Agency's Defense Data Network (DDN).

Key Words: Exponential packet transit time; number of packets per message; receiver blocking; infinite server queue; Erlang loss formula.

1. INTRODUCTION

The purpose of this report is to introduce an initial model for the effect of message size upon data transfer time in the Defense Communications Agency's Defense Data Network (DDN). The model is quite simple, but does contain certain qualitative features of the DDN, as we currently understand them. It is hoped that by explicitly formulating and exploring such a model we can elicit comments that will lead to improvement, and eventually result in a useful tool for some aspects of system planning.

Our modeling approach is to focus on the experience of a single *tagged job*, J, that is to be accomplished in an environment of many other jobs all utilizing, and competing for, a number of resources: packet switches and the links interconnecting them. We think of a job as meaning the transfer of a certain number of packets (of data) from one node to another in packet switching fashion. We assume to start with that the aspects of behavior of the particular tagged job, J, that are environmentally influenced are not changed by changes made in the way J's transmission is managed: in particular, attempts to transmit J in one piece, or in minimal packet-sized, "messages," or at some intermediate degree of message size (m packets) does not influence the environment. At a later stage of analysis we will attempt to represent the "environmental impact" of a policy that sends all (or many) messages of about the same size; perhaps this can be done by a process of iteration. Our initial approach seems necessitated by the complexity of the true packet switching setup, which involves buffers (queues) at the various nodes (switches) having limited capacities and elaborate protocols for traffic management. We start in as simple a way as seems qualitatively correct, and then plan to ratchet the model up to deal with whatever issues come to light.

With this preamble, here is our initial message transfer protocol and model formulation.

At an initial instant (t=0) a group of $m \geq 1$ packets is to be sent from a Source buffer, S, to a destination buffer, D, with the following protocol in effect:

Protocol.

a) The m packets are sent out simultaneously via the network, destined for D. The first packet to arrive and find space for the m-packet message occupies a one-packet space and reserves space for m−1 others, and sends an acknowledgement back to S.

b) If no acknowledgement is received by S in time δ each packet is re-transmitted; this action is repeated until acknowledgement occurs. Such is necessary because packets may encounter full buffers in transit and be lost, or may reach D, encounter a too-full buffer and be lost, or be lost for other reasons.

c) Once the initial reservation is made a time must elapse until the remaining m−1 packets (originals, or duplicates thereof) reach D: each subsequent packet experiences delays similar to the first, but need not reserve space in D's buffer.

Next we construct a mathematical model that reflects the effect of the above protocol and behavior.

2. PROBABILITY MODEL FOR MESSAGE SERVICE TIME: TIME FOR THE FIRST ENTRY TO RESERVE BUFFER.

Packets transmitted from S to D on a network may travel by different routes and arrive somewhat independently, at which moment there is an attempt to capture space in a finite buffer owned by D. The fact that a particular message's packet must compete with others suggests these convenient assumptions:

A_1: transit times T of individual packets from the tagged message are independently and identically distributed according to F(t); the form of this latter distribution will be discussed later;

A_2: when a packet of the tagged message arrives at D's buffer then if it is the first and is to be successful it must capture a space of size m—space for m–

1 packets besides itself. We assume that the probability of such capture, denoted by a=a(m), may be treated as a success probability in independent (Bernoulli) trials. This is suggested if each packet of the tagged message arrives independently and in the general company of others from many other sources arriving at, and demanding space in, D's buffer. Actual determination of the probability a(m) will be discussed later.

On the basis of the Protocol and assumptions $\mathbf{A_1}$ and $\mathbf{A_2}$ it is now possible to write down an expression for the probability that T(1), the time of first packet entry to D's buffer, exceeds t:

$$P\{T(1)>t\} = \left(\prod_{i=0}^{[t/\delta]} (a\bar{F}(t-i\delta)+\bar{a}) \right)^m$$

$$= \left(\prod_{i=0}^{[t/\delta]} (1-aF(t-i\delta)) \right)^m . \qquad (2.1)$$

As usual $[t/\delta]$ is the largest integer $\leq t/\delta$.

This simply expresses the fact that no packet—either an original or a copy—has entered D's buffer by time t. Given F and a, and also δ and m, one can evaluate the above probability, but in general this is a tedious numerical task.

3. APPROXIMATION TO THE DISTRIBUTION OF T(1), THE RESERVATION TIME

Note that taking logarithms of both sides of (2.1) gives

$$\ln P\{T(1)>t\} = m\left(\sum_{i=0}^{[t/\delta]} \ln(1-aF(t-i\delta))\right)$$

$$= \frac{m}{\delta}\left(\sum_{i=0}^{[t/\delta]} \ln(1-aF(t-i\delta))\right)\delta \tag{3.1,a}$$

$$\sim \frac{m}{\delta}\int_0^t \ln(1-aF(x))\,dx \tag{3.1,b}$$

where we have viewed the sum as a Riemann sum approximating the integral. Use of the Euler-MacLaurier sum formula may produce useful improvements.

To get further explicit results assume that $F(x) \sim (\lambda x)^\beta$; $\lambda, \beta>0$ as $\lambda \to 0$. This is a property of the Weibull and Gamma distributions, both reasonable candidates for describing transit times. Putting this into (3.1,b) yields after further Taylor expansion

$$\ln P\{T(1)>t\} \sim \frac{-ma}{\delta}\int_0^t (\lambda x)^\beta\,dx = -\left(\frac{ma}{\delta}\lambda^\beta\right)\frac{t^{\beta+1}}{\beta+1}$$

so

$$P\{T(1)>t\} \sim \exp\left[-\frac{ma\lambda^\beta}{\delta}\cdot\frac{t^{\beta+1}}{\beta+1}\right] \tag{3.2}$$

and the distribution of T(1) is seen to be approximately Weibull. If $\beta=1$, the transit times T are exponentially distributed and the distribution of T(1) is approximately Rayleigh. In this case,

$$E[T(1)] = \int_0^\infty P\{T(1)>t\}dt \sim \int_0^\infty \exp\left\{\frac{-ma\lambda}{2\delta}x^2\right\}dx = \sqrt{\frac{\pi}{2}}\sqrt{\frac{\delta}{\lambda ma}} \qquad (3.3)$$

as explicit a formula as is likely to be available. General Weibull ($\beta>0$ arbitrary) results can be found in terms of gamma functions. Note that the appearance of the square root in (3.3) suggests that precision of determination of λ and particularly a(m) may not be important.

4. THE D-BUFFER MODEL

The time until a reservation is made at D's buffer is obviously affected by the size of that buffer. Suppose it has capacity B, in packets. Then it is reasonable to make the following initial model, the inclusion of which into our previous model we term

A_3: the long-run or steady-state probability distribution of the contents of the D-buffer is

$$\pi_j = K(B)\,\alpha^j/j! \quad j = 0, 1, \ldots B \qquad (4.1)$$

where $\sum_{j=0}^{B} \pi_j = 1$ determines K(B). In other words, we characterize the buffer state as being truncated Poisson, as is true of an M/G/B/B queueing system in equilibrium. Of course the parameter α must be estimated, and depends upon the rate at which new messages arrive at the particular buffer as well as their "holding time" in that buffer. The buffer size, B, is a decision variable.

Now given the above we let

a(m) ≡ Probability that space for at least m packets exists in the buffer at a "random" time

$$= \sum_{j=0}^{B-m} \pi_j \tag{4.2}$$

with π_j being given by (4.1). Furthermore, a(m) will be used as before, i.e., as the success probability in a Bernoulli trial (coin-flip) scheme.

5. THE TIME TO COMPLETE MESSAGE TRANSMISSION

Having made a reservation at D's buffer with the arrival of the first packet to find $\geq m$ packet slots available we now wish to compute (an approximation to) the expected time to fill the remaining m-1 slots. It may be seen that the number N_j of duplicate packets outstanding for packet j is approximately Poisson with mean $1/(\lambda\delta)$; simply view the network transit time, here taken to be exp(λ), as a service time in an M/M/∞ queueing process with arrival rate δ, and consider the steady-state situation. Now let S_j be the time until outstanding packet j enters the buffer. Then if we use the Poisson approximation for re-transmits (See Appendix B) we find

$$P\{S_j > s\} = E\left[e^{-\lambda s N_j} e^{-(s/\delta)\left[1-\left(1-e^{-\lambda s}\right)/\lambda s\right]}\right]$$

$$= e^{-s/\delta}, s > o, \tag{5.1}$$

where we have used the fact that $E\left[e^{-\lambda s N_j}\right] = e^{-1/\lambda\delta\left[1-e^{-\lambda s}\right]}$. Now assuming that all subsequent transits, S_j, j = 1, 2, ... m−1 are independently and exponentially distributed according to (5.1) it turns out that the expected time to complete filling the m−1 reserved slots is distributed as

$$P\left\{T(2) \equiv \max S_j \leq s\right\} = \left(1-e^{-s/\delta}\right)^{m-1}. \tag{5.2}$$

Consequently

$$E[T(2)] = E[\text{Max } S_j] = \delta\left[1 + \frac{1}{2} + \ldots + \frac{1}{m-1}\right] \sim \delta \ln(m-1). \tag{5.3}$$

Thus the total time to transfer an m-packet message is $T_m = T(1) + T(2)$, with expectation

$$E[T_m] \equiv E[T(1)] + E[T(2)]. \tag{5.4}$$

6. OPTIMIZING THE MESSAGE SIZE

If a large number of packets, i.e., making up a data base, is to be transferred in messages of size m then renewal theory, see Ross (1985), can be invoked to show that the long-run rate of transfer of data packets is of the form

$$r(m) = \frac{m}{E[T_m] + c} \tag{6.1}$$

where c is the time to acknowledge a successful transfer. In words, a *reward* of m packets is received every time a *cycle* of length $E[T_m] + c$ completes. It is intuitively apparent that an optimal value of m will often exist. It can be found numerically, as indicated in the following simple and illustrative (but doubtless unrealistic) example.

In this example a message size of 32 packets maximizes the transfer rate. A larger size is apparently much less effective because of the difficulty of reserving space in D's buffer. Presumably a smaller size is less effective because acknowledgements consume too much time.

EXAMPLE

$\alpha = 25,\ B = 50$

$\beta = 1 = \lambda,\ \delta = \frac{1}{2},\ c = 1$

m (# packets in message)	r(m) (Rate of transfer)	E[T(1)]	E[T(2)]
2	0.94	0.63	0.50
8	3.07	0.31	1.30
14	4.95	0.24	1.59
20	6.70	0.21	1.77
26	8.22	0.25	1.91
32	9.06	0.52	2.01
38	6.70	2.56	2.10

Figure 1 displays the packet transmission rate as a function of the number of packets in a message for two values of the retransmission time δ. The other parameters are $\lambda = 0.5$, $B = 50$, $c = 1$, $\alpha = 0.5$. Note that the transmission rate for $\delta = 0.75$ is always less than that for $\delta = 0.5$. This behavior is due to the fact that the model for D's buffer is not affected by the value of δ. However, since all sources would be following the same retransmission policy a decrease in δ may increase the congestion at D's buffer.

Figure 2 displays the packet transmission rate for a model in which D's buffer is affected by δ. D's buffer is still modeled by an M/M/B/B queue. However, the parameter of the limiting distribution of the queue is $\alpha = 12.5/\delta$ rather than $\alpha = 25$ as in Figure 1. The parameter δ takes 3 values, $\delta = 0.4$, $\delta = 0.5$ and $\delta = 0.7$. The other parameters are as in Figure 1. Note that $\delta = 0.5$ has the largest maximizing transmission rate of the three. If δ is too large, too much time is spent waiting to recover from a last packet. If δ is too small, D's buffer is more likely to be full and packets from S will be lost more frequently. The parameterization of α is ad-hoc for this model. Future models will address the effect of other sources' retransmissions on the environment a tagged packet is subjected to.

7. CONCLUSIONS

We have presented an initial model that illustrates the interplay of various system factors influencing choice of optimal message size in data transfer. This study is only a beginning, and must be expanded and improved in order to be truly useful. Nevertheless it does capture some features of the real situation.

Examination of the table suggests that there is an advantage to increasing the message size m initially: this hastens the occurrence of a reservation, i.e., shortens E[T(1)]. However, too large an m quickly makes reservations in the D-buffer extremely difficult and extends the time E[T(1)] greatly. Naturally, beneficial changes in both δ and B can influence the transfer rate as well; an overall optimization attempt should be placed in the agenda. Finally the effect of changing the retransmission time δ on the environment of other packets that a tagged packet encounters is another subject of future study.

APPENDIX B. POISSON APPROXIMATION TO THE DISTRIBUTION OF THE ARRIVAL TIME TO THE BUFFER BY AN OUTSTANDING PACKET AFTER A MESSAGE RESERVATION HAS BEEN MADE

Suppose a packet of a message has just made a reservation at D's buffer. We are interested in an approximation to the distribution of the additional time required for the j^{th} outstanding packet of the message to arrive at D. We will assume the network transit times for packets are iid exponential with mean λ^{-1}.

We will approximate the process of retransmission of packet j by a Poisson process with rate $\frac{1}{\delta}$ independent of the other packets. Although packets are retransmitted at constant intervals, δ', this approximation may be justified if many of the packets are lost or damaged, in this case $\delta > \delta'$.

Let N_j be the number of outstanding duplicates for packet j at the time the reservation is made. We will approximate the distribution of N_j by the limiting distribution of an $M/M/\infty$ queue with arrival rate $\frac{1}{\delta}$ and mean service time $\frac{1}{\lambda}$. This approximation leads to N_j having a Poisson distribution with mean $(\lambda\delta)^{-1}$.

Let U_j be the smallest arrival time at D of the additional duplicate packets sent after the reservation is made.

$$P\{\mathrm{U_j} > s\} = \sum_{n=0}^{\infty} \exp\left\{-\frac{1}{\delta}s\right\} \frac{\left[\frac{1}{\delta}s\right]^n}{n!} \left[\frac{1}{s}\int_0^s \exp\{-\lambda(s-u)\}du\right]^n$$

$$= \exp\left\{-\frac{1}{\delta}s\left[1-\left(1-e^{\lambda s}\right)/\lambda s\right]\right\}$$

$$= \exp\left\{-\frac{1}{\lambda\delta}\left[\lambda s-\left(1-e^{-\lambda s}\right)\right]\right\}.$$

Let $\mathrm{S_j}$ be the time of arrival at D of the first of the duplicate packets for outstanding packet j.

$$P\{\mathrm{S_j} > s\} = E\left[\left(e^{-\lambda s}\right)^{N_j}\right] P\{\mathrm{U_j} > s\}$$

$$= \exp\left\{-\frac{1}{\delta\lambda}\left[1-e^{-\lambda s}\right]\right\} \exp\left\{-\frac{1}{\delta\lambda}\left[\lambda s-\left(1-e^{-\lambda s}\right)\right]\right\}$$

$$= \exp\left\{-\frac{1}{\delta}s\right\}.$$

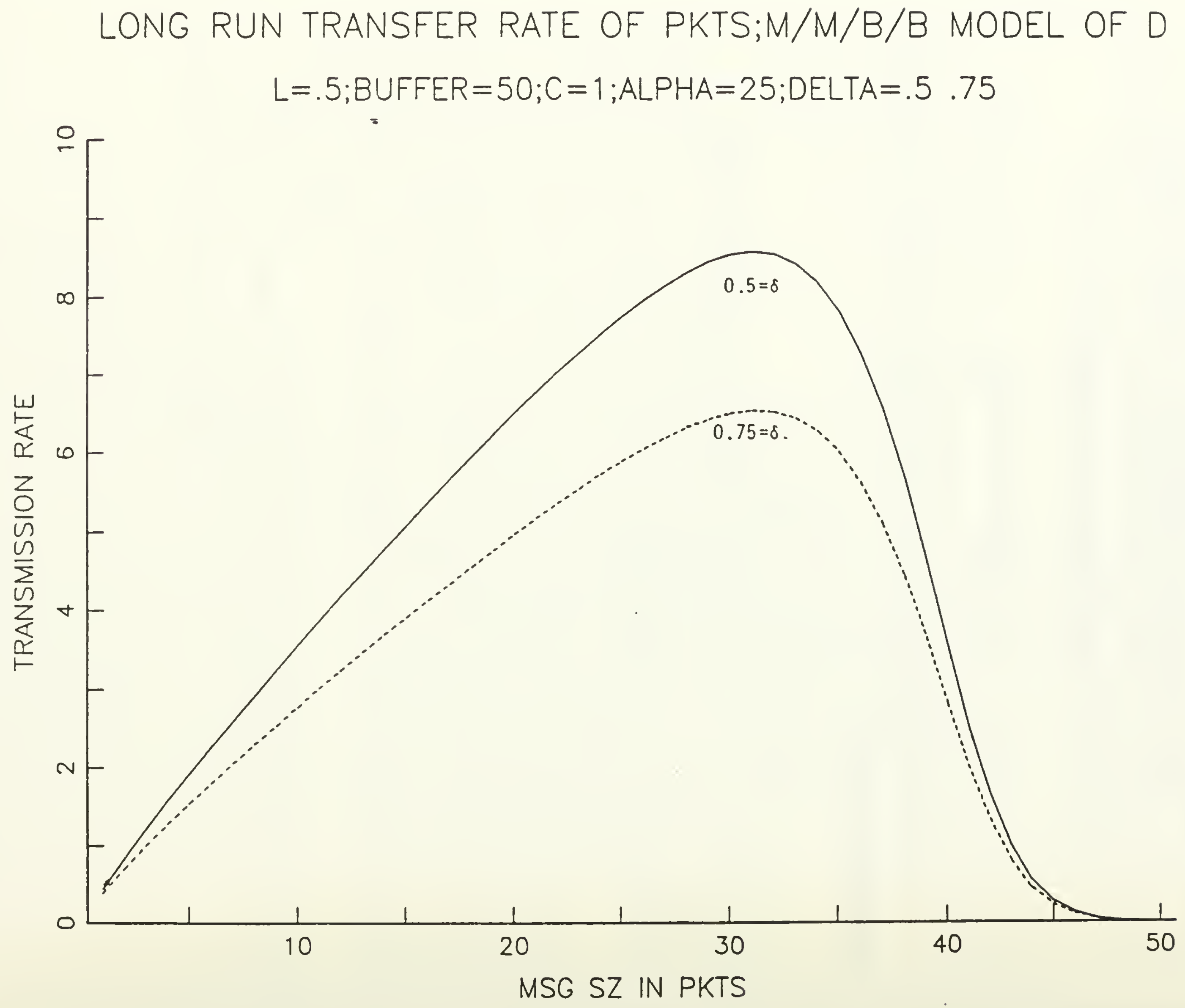

LONG RUN TRANSFER RATE OF PKTS;M/M/B/B MODEL OF D
L=.5;BUFFER=50;C=1;ALPHA=25;DELTA=.5 .75
TRANSMISSION RATE
0
2
4
6
8
10
0.5=δ
0.75=δ.
10
20
30
40
50
MSG SZ IN PKTS

Figure 1

LONG RUN RATE OF TRANSFER OF DATA PKTS;M/M/B/B MODEL FOR D

L=.5;BUFFER=50;C=1;ALPHA=12.5÷DELTA;DELTA=.4,.5,.7

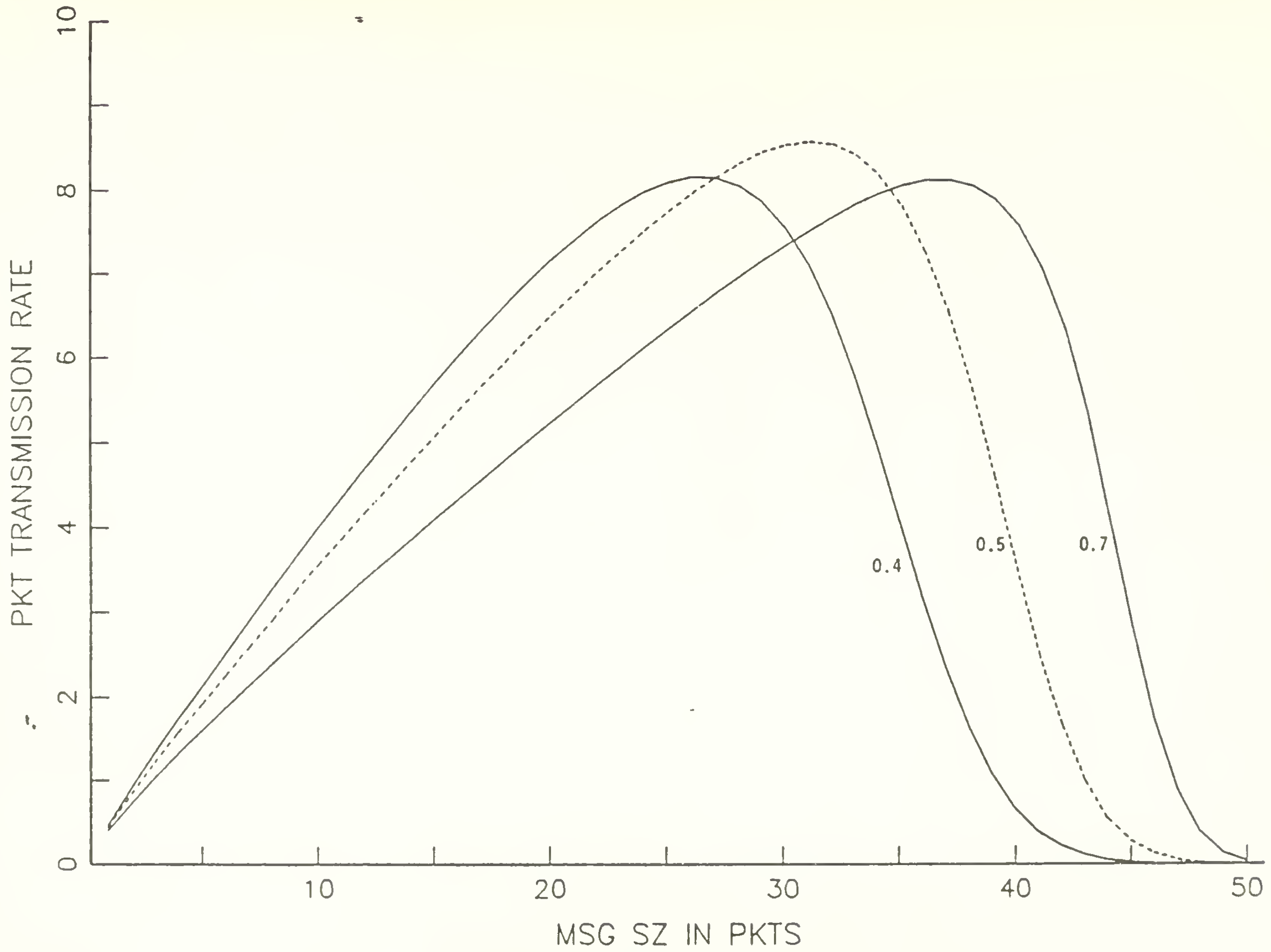


Figure 2

REFERENCES

BBN Communication Corporation. *Computer Program Functional Specification Packet Switching Node (PSU).* BBN Report No. 6874, Document No. B006, 15 July 1988.

S. M. Ross. *Introduction to Probability Models.* Third Edition, Academic Press, New York, 1985.

DISTRIBUTION LIST

	NO. OF COPIES
Library (Code 0142) Naval Postgraduate School Monterey, CA 93943-5000	2
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Office of Research Administration (Code 012) Naval Postgraduate School Monterey, CA 93943-5000	1
Library (Code 55) Naval Postgraduate School Monterey, CA 93943-5000	1
Operations Research Center, Rm. E40-164 Massachusetts Institute of Technology Attn: R. C. Larson and J. F. Shapiro Cambridge, MA 02139	1
Koh Peng Kong OA Branch, DSO Ministry of Defense Blk 29 Middlesex Road SINGAPORE 1024	1
Arthur P. Hurter, Jr. Professor and Chairman Dept of Industrial Engineering and Management Sciences Northwestern University Evanston, IL 60201-9990	1
Institute for Defense Analysis 1800 North Beauregard Alexandria, VA 22311	1
Professor H. G. Daellenbach Dept of Operations Research University of Canterbury Christchurch, NEW ZEALAND	1
Dept of Operations Research (Code 55) Naval Postgraduate School Monterey, CA 93943-5000	93